

181 CAGATGCTGGATGACCCAGAGGAGCATTTGAAGTTCCAAAGAGCTCTGGGAAAGTG 240
241 CTGCTGGATGCAATTTGGCAACCACTTCAACCTCTGTGAAAGGTGACTATTTGGCTCGAG 300
241 CTGCTGGATGCAATTTGGCAACCACTTCAACCTCTGTGAAAGGTGACTATTTGGCTCGAG 300
301 TTTCTGATCACAAAAGATCAGGTCGTGGCTGATCTCTTAAACCAATTTGAAACAG 360
301 TTTCTGATCACAAAAGATCAGGTCGTGGCTGATCTCTTAAACCAATTTGAAACAG 360
361 ATTAGAAGCCAAAGACGTTGTTGTAAGTTGTGGTAAATTTCTTCGCTGACAC 420
361 ATTAGAAGCCAAAGACGTTGTTGTAAGTTGTGGTAAATTTCTTCGCTGACAC 420
421 ACACAACTCCAAAGAAATCTCACAAAGTACTGTTCGGCTGAGAGTGAAGCAGACTTG 480
421 ACACAACTCCAAAGAAATCTCACAAAGTACTGTTCGGCTGAGAGTGAAGCAGACTTG 480
481 GCTCAAGGAGGTTGACGTGTAAATGACACAGCGAGCTCTCTGATTTCAACATTTG 540
481 GCTCAAGGAGGTTGACGTGTAAATGACACAGCGAGCTCTCTGATTTCAACATTTG 540
541 CAATCTGAGATTGGGATTTTGAATGAAAGCTTGGACAGAGACCTTACCAAAAATAA 600
541 CAATCTGAGATTGGGATTTTGAATGAAAGCTTGGACAGAGACCTTACCAAAAATAA 600
601 TACATCTCTCAGAGAACCACTAGAGACCAAAATGCTGAAATTTTCAACATAACCAATT 660
601 TACATCTCTCAGAGAACCACTAGAGACCAAAATGCTGAAATTTTCAACATAACCAATT 660
661 GGAACAACACAGCAGAAATCAGATTTCCAGCTCTAGAGATTGCCGTGGCTAGAGATG 720
661 GGAACAACACAGCAGAAATCAGATTTCCAGCTCTAGAGATTGCCGTGGCTAGAGATG 720
721 TATGAATCCGGTTGACCCGCGCAAGAGCAGGGAAGGCAAGAGATCAATCTGCGCTT 780
721 TATGAATCCGGTTGACCCGCGCAAGAGCAGGGAAGGCAAGAGATCAATCTGCGCTT 780
781 GCCAACAAGGAAATTTAGTGTTCAGGGTTTCACTAAGATCAATGCTTCAACTGAGGCG 840
781 GCCAACAAGGAAATTTAGTGTTCAGGGTTTCACTAAGATCAATGCTTCAACTGAGGCG 840
841 AAGGTCGGAAGCTGAGCTTCAAGAGGAAGCGCTTCTCATCAAGCTCCGGCCAGATGCC 900
841 AAGGTCGGAAGCTGAGCTTCAAGAGGAAGCGCTTCTCATCAAGCTCCGGCCAGATGCC 900
901 AATAGTCGTACAGGATACCTTGAATTCCTGATGCGCAATCTTGAAGGCCAAACCA 960
901 AATAGTCGTACAGGATACCTTGAATTCCTGATGCGCAATCTTGAAGGCCAAACCA 960
961 TTTCTGAAAATCTGTGTGAACATCATGCTCTTGAACCTTTTGAAGGCCAAACCA 1020
961 TTTCTGAAAATCTGTGTGAACATCATGCTCTTGAACCTTTTGAAGGCCAAACCA 1020
1021 AAGCCCAAGCCGCTCTCTTAAAGCCGAGGATCATCTTCTGATTCAGTGTGCACTCAG 1080
1021 AAGCCCAAGCCGCTCTCTTAAAGCCGAGGATCATCTTCTGATTCAGTGTGCACTCAG 1080
1081 AAGCAGGTTCTGACATGTAAAGAGAGCATTAAGAGGTGCAATTTTGAAGGAAG 1140
1081 AAGCAGGTTCTGACATGTAAAGAGAGCATTAAGAGGTGCAATTTTGAAGGAAG 1140
1141 CACAGAAATTCATCTATCCGAGACCTTGTTCACAGCTTACAAATCTGAATTTGGAA 1200
1141 CACAGAAATTCATCTATCCGAGACCTTGTTCACAGCTTACAAATCTGAATTTGGAA 1200
1201 GTGCTGAGACAGTCTCAGAGAGACACAGCTTACATTTGAGAGAGGTGCCGATCTCA 1260
1201 GTGCTGAGACAGTCTCAGAGAGACACAGCTTACATTTGAGAGAGGTGCCGATCTCA 1260
1261 GGGGAGCAGAGCTGCGGCGAGGAAAGAACGAGAGTTTCGCGCGGAGACCGGAGTCTG 1320
1261 GGGGAGCAGAGCTGCGGCGAGGAAAGAACGAGAGTTTCGCGCGGAGACCGGAGTCTG 1320

1261 GGGGAGCAGAGCTGCGGCGAGGAAAGAACGAGAGTTTCGCGCGGAGACCGGAGTCTG 1320
1321 CACCGAAGCCCTGCGGCGAGAGAGAGCCCGGGGTAAACAAGCAGCGGAGCGGACCGCG 1380
1321 CACCGAAGCCCTGCGGCGAGAGAGAGCCCGGGGTAAACAAGCAGCGGAGCGGACCGCG 1380
1381 TCGGCGCCACGAGAGAGAGAGAGGTCTTTAAGATAGAGACCAGCAGATTAACCT 1440
1381 TCGGCGCCACGAGAGAGAGAGAGGTCTTTAAGATAGAGACCAGCAGATTAACCT 1440
1441 CAGCCCCGAGCCAGAGCAGAGCTCTCTGATCAGCGGCTGCACTCTTCCGAGCTGCT 1500
1441 CAGCCCCGAGCCAGAGCAGAGCTCTCTGATCAGCGGCTGCACTCTTCCGAGCTGCT 1500
1501 GTGAAGCTGAGGGGGAGGTGGCCCTGCAACGATCCTTGTCCCAACTGAGCCCC 1560
1501 GTGAAGCTGAGGGGGAGGTGGCCCTGCAACGATCCTTGTCCCAACTGAGCCCC 1560
1561 GACACCAAGCAGGCTCTCTCTGATCAGCGGCTGCACTCTTCCGAGCTGCTG 1620
1561 GACACCAAGCAGGCTCTCTCTGATCAGCGGCTGCACTCTTCCGAGCTGCTG 1620
1621 ACGAGCATGAGATGAGGCGCGAGAGAGATTCCTGATTAAGCTGATTAAGCTTCAAT 1680
1621 ACGAGCATGAGATGAGGCGCGAGAGAGATTCCTGATTAAGCTGATTAAGCTTCAAT 1680
1681 GCTAAGAAAGTGTCTAACCAAGAGCAATATCTGAAGATCTGAAAGTTTCACTTGG 1740
1681 GCTAAGAAAGTGTCTAACCAAGAGCAATATCTGAAGATCTGAAAGTTTCACTTGG 1740
1741 TGGTTTCAGAGCAGTGAAGCAAGAGACGCGATGCGGAAAGCACTGAAAAGTCTCAT 1800
1741 TGGTTTCAGAGCAGTGAAGCAAGAGACGCGATGCGGAAAGCACTGAAAAGTCTCAT 1800
1801 TTCCCGAATTTTGAACCTTTTGAACAAATTTTCACTAATTTTCTCAAGAAATTAAGCA 1860
1801 TTCCCGAATTTTGAACCTTTTGAACAAATTTTCACTAATTTTCTCAAGAAATTAAGCA 1860
1861 CGACTTGCCTGTGGAGAGGCGCTCAATGCGCAATCAGAGATTAACAAGAAATGCGC 1920
1861 CGACTTGCCTGTGGAGAGGCGCTCAATGCGCAATCAGAGATTAACAAGAAATGCGC 1920
1921 GATGTCATGCTGAAGAACATTCAGGGCATGAAGCACTTGGCGCTCACCTGTGAAGCAC 1980
1921 GATGTCATGCTGAAGAACATTCAGGGCATGAAGCACTTGGCGCTCACCTGTGAAGCAC 1980
1981 AAGCAGGCTTGAAGGCGCTTGAAGATGAAGATCAAGAGCTCCCGGCGCTGAAGAACTTC 2040
1981 AAGCAGGCTTGAAGGCGCTTGAAGATGAAGATCAAGAGCTCCCGGCGCTGAAGAACTTC 2040
2041 TGCAGAGCTTTGAGCTCAGAGAGGTGTGTTACTTACCGCTCAACACTTCTCTGCGG 2100
2041 TGCAGAGCTTTGAGCTCAGAGAGGTGTGTTACTTACCGCTCAACACTTCTCTGCGG 2100
2101 CCACTGACCGGCTCATGCACTTACAGAGAGTCTTGAAGCGGCTGTGCAACACACCCG 2160
2101 CCACTGACCGGCTCATGCACTTACAGAGAGTCTTGAAGCGGCTGTGCAACACACCCG 2160
2161 CCGAGCCACGCGAATTTAGAGGACTGCGGAGCCGCTTTTGGAGAGATCAAGAGATGAGT 2220
2161 CCGAGCCACGCGAATTTAGAGGACTGCGGAGCCGCTTTTGGAGAGATCAAGAGATGAGT 2220
2221 GCAAGCTCAGCGGATGATCAAGATGAGAAATTTCCAGAGCTGCAAGACTCAAG 2280
2221 GCAAGCTCAGCGGATGATCAAGATGAGAAATTTCCAGAGCTGCAAGACTCAAG 2280
2281 AAAAATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 2340
2281 AAAAATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 2340
2341 AGCTTCAGCAAGCTTCCGGGAGAGGGGCTCAGAGAGCGCATTTTCTTCTGTTCAAGAC 2400
2341 AGCTTCAGCAAGCTTCCGGGAGAGGGGCTCAGAGAGCGCATTTTCTTCTGTTCAAGAC 2400
2441 AGCTTCAGCAAGCTTCCGGGAGAGGGGCTCAGAGAGCGCATTTTCTTCTGTTCAAGAC 2400
2441 AGCTTCAGCAAGCTTCCGGGAGAGGGGCTCAGAGAGCGCATTTTCTTCTGTTCAAGAC 2400


```

QY 2401 GTCCTGATATACGAGCGCGGGGCTGACGGCCCTCAATGATTTAAAGTCCACGGGCGAG 2460
DB 2401 GTCTCTGATATACGAGCCCGGGGCTGACGGCCCTCAATGATTTAAAGTCCACGGGCGAG 2460
QY 2461 CTCCTCCCTATATGAGCATGAGATTGAGAGAGCGAAGACGAGTGGGGGGTGTCCCACTGC 2520
DB 2461 CTCCTCCCTATATGAGCATGAGATTGAGAGAGCGAAGACGAGTGGGGGGTGTCCCACTGC 2520
QY 2521 CTGACCTCTCGGGGGCCGAGCGGCGAGTCCATCATGTGGCCCGCACTTCTCGGTCCGAGATG 2580
DB 2521 CTGACCTCTCGGGGGCCGAGCGGCGAGTCCATCATGTGGCCCGCACTTCTCGGTCCGAGATG 2580
QY 2581 GAGAAGGGGTTGAGGACATCCAGTGGGCAATTGACCTGGCGGAGAAAGACAGACGGCC 2640
DB 2581 GAGAAGGGGTTGAGGACATCCAGTGGGCAATTGACCTGGCGGAGAAAGACAGACGGCC 2640
QY 2641 GCCCTGAGTCTCTGGCGAGACAGCCCTGACAAACAGTCCCTGATGAGAGCCAGCGCG 2700
DB 2641 GCCCTGAGTCTCTGGCGAGACAGCCCTGACAAACAGTCCCTGATGAGAGCCAGCGCG 2700
QY 2701 GCTGACCAAGAGTCAAGAGATGACCTGAGCGCTCCGCGACATGCTGAGAGCCAGGCC 2760
DB 2701 GCTGACCAAGAGTCAAGAGATGACCTGAGCGCTCCGCGACATGCTGAGAGCCAGGCC 2760
QY 2761 CGGACCGGGGGCAACAAATGAGCACTGTGCTGGGACCGGAAACCGAGCGTCTCCATG 2820
DB 2761 CGGACCGGGGGCAACAAATGAGCACTGTGCTGGGACCGGAAACCGAGCGTCTCCATG 2820
QY 2821 GTGAGACTTGACATCCGAGTGGAGAAATCAGTTCTGAGAAACCTGCTGAGAAATTCAA 2880
DB 2821 GTGAGACTTGACATCCGAGTGGAGAAATCAGTTCTGAGAAACCTGCTGAGAAATTCAA 2880
QY 2881 AACAGCAAGGGGTGGGCAAGAGTGTGGGTGTTCACAACTTCTGCTGTCTTCTTAC 2940
DB 2881 AACAGCAAGGGGTGGGCAAGAGTGTGGGTGTTCACAACTTCTGCTGTCTTCTTAC 2940
QY 2941 AATGACACAGGACAAATATCCCTTGGCAGCTGCTGTGCTGGGCTATCTGCTGAC 3000
DB 2941 AATGACACAGGACAAATATCCCTTGGCAGCTGCTGTGCTGGGCTATCTGCTGAC 3000
QY 3001 ATCCCTCTGAGTCCGAGAAACATCCAGAAAGATCAGTGTTCAGAGTGCATTCAGATCC 3060
DB 3001 ATCCCTCTGAGTCCGAGAAACATCCAGAAAGATCAGTGTTCAGAGTGCATTCAGATCC 3060
QY 3061 CACGCTCTACTTCAAGGCGGGAACGAGTACACGTTGGAAGGTGATGAGATGATC 3120
DB 3061 CACGCTCTACTTCAAGGCGGGAACGAGTACACGTTGGAAGGTGATGAGATGATC 3120
QY 3121 CGCAGTGCACACGCTCTGCTCGGACCCCAAGTGTGAGCCACAAAGATCTCTTGTG 3180
DB 3121 CGCAGTGCACACGCTCTGCTCGGACCCCAAGTGTGAGCCACAAAGATCTCTTGTG 3180
QY 3181 TATTGATGCGCGGACACACTCGTTTCGCGAGTGGTCTTCTCGAGAGACGTTTCTT 3240
DB 3181 TATTGATGCGCGGACACACTCGTTTCGCGAGTGGTCTTCTCGAGAGACGTTTCTT 3240
QY 3241 GTTCTGTATTTAATGAAGCTGTGTAATTAACCTGTCTGAAATCAAAAACATGGCTT 3300
DB 3241 GTTCTGTATTTAATGAAGCTGTGTAATTAACCTGTCTGAAATCAAAAACATGGCTT 3300
QY 3301 CCAGAGAGCTCTGCTGCTCCACAGCGGGTTTTTAAACCCGACCTCTCAGCGTTTGA 3360
DB 3301 CCAGAGAGCTCTGCTGCTCCACAGCGGGTTTTTAAACCCGACCTCTCAGCGTTTGA 3360
QY 3361 TGAACAGCGCTCCACCTCAGTCTGAGATCCGCTGGGGGCGCTTCTTTAGCTAGTG 3420
DB 3361 TGAACAGCGCTCCACCTCAGTCTGAGATCCGCTGGGGGCGCTTCTTTAGCTAGTG 3420
QY 3421 CCAAGTATTTAAACATTTGCTATT 3442
DB 3421 CCAAGTATTTAAACATTTGCTATT 3442

```

```

RESULT 2
AB008430 3442 bp mRNA linear PRI 13-FEB-1999
LOCUS
DEFINITION Homo sapiens mRNA for CDEP, complete cds.
ACCESSION AB008430
VERSION AB008430.1 GI:2766164
KEYWORDS CDEP.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.
REFERENCE 1 (bases 1 to 3442)
AUTHORS Koyano,Y., Kawamoto,T., Shen,M., Yan,W., Noohiro,M., Fujii,K. and
Kato,Y.
TITLE Molecular cloning and characterization of CDEP, a novel human
protein containing the ezrin-like domain of the band 4.1
superfamily and the Dbl homology domain of Rho guanine nucleotide
exchange factors
JOURNAL Biochem. Biophys. Res. Commun. 241 (2), 369-375 (1997)
PUBMED 9425278
REFERENCE 2 (bases 1 to 3442)
AUTHORS Koyano,Y., Kawamoto,T. and Kato,Y.
JOURNAL Direct Submission
TITLE Submitted (22-OCT-1997) Takeshi Kawamoto, Hiroshima University
School of Dentistry, Department of Biochemistry, 1-2-3 Kasumi
Minami-Ku, Hiroshima, Hiroshima 734, Japan
(E-mail:tkawamo@ipc.hiroshima-u.ac.jp, Tel:082-257-5688,
Fax:082-257-5629)
FEATURES
source Location/Qualifiers
1..3442
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/cell_type="chondrocyte"
/tissue_type="cartilage"
/dev_stage="embryo"
49..3186
/function="Rho Guanine Nucleotide Exchange Factor"
/notes="Band 4.1 superfamily"
/codon_start=1
/product="CDEP"
/protein_id="BAA24267.1"
/db_xref="GI:2766165"
/translation="MGEIEQRPTPGSRILGAPENSGISTLERGQKPPPTSGKLVSIKI
QMLDDTOEALFEVPOBAPGKVLDAVONHNLVEGDFGLEPDHKKITVMDLILPVI
KOIRBPKHVIVFVVKFPDPDHOEELTRYLPMQVQEDLAQGLTNDPSAALI
SHIVOSRIGDFDEALDREHLAKNKYIPODDALEKIVFHHNHIGOTPAESPFOLEI
ARLEMYGRIHPADREGTKYNLVANTGILVPGAPFKYNAFMNAKPKKLSFKRRF
LTKLRPDANSATQDTLEFLMARSDFCKSFVKI CVENHAFRLFEERPKRKYLSRG
SSFRSGRTQKQVLDYVREGHKVQFERKSHKHSISLASQPELVSEVLSQOS
TSVTFGEABSPGSCSRGKEPKVSAGBPSHPSPAPRSPAGKQNDGASAPTEE
BEVVKDRTQGSQKPPQPSSTLSGSHLSLVNSQGVVPAVNTLSPNLSPTKQ
ASPLISPLINDACRTDDEDEGRKRPPTDPAFLYLAERVSTTERTYKDLBITWFMF
OSTVSKEDAMPKSLIAPNPEPLKHKHTNFKETIABRIALMBERSNAOIPDYRIG
DMLKNTQGMKHLAHMKHSBALENGKISRLNEFCDFELQKCYIPLNTFL
LRLRLHMYKQVLERLCKHPPSHADRDCAALAEITEVAQDLHGTIKENRQKL
HELKLDLIGIDLIVVGRGFRILGSLKSGGLQORMFLEFNDVLLYTSRLTYSNQ
FVHGQQLPLYGMITIESEDEMGVPCILTLRGOROSIIVAASGRSMEKVEDIOAID
LAEKSSPAPLEPLASPDNKSPPDAETADODESDLSASRTSLERQAPRHGNTAVH
CMHRVTSVMDPSTFAVENOLSGNLLRKRKNSGMQKLMVPTNCLFPYKSHQNHNP
LASLPLIGSLTIPESENITQDVFYFKLHFSSHVYTPRAESBETTERMEVIRSATSS
ASRPHVLHKSLELV"
3442
/notes="50 a nucleotides"

```

polyA_site

ORIGIN

Query Match 100.0%; Score 3442; DB 5; Length 3442;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3442; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCCGACGCGCGCGCTGTGAGATATTTCTAAGCGGCTTTCATGATGAGAGAAATA 60

|||||
Db 1 CGCCGAGCCGCGCGCTGTGGAGATATCTCTAAGCCGCTTTCATCATGGAGAAATA 60
QY 61 GAGCAGAGGCCAGCCCCCAGATATCAAGATCGGGGGCCCCGGAAAAATTCGGGGATCACTAC 120
Db 61 GAGCAGAGGCCAGCCCCCAGATATCAAGATCGGGGGCCCCGGAAAAATTCGGGGATCACTAC 120
QY 121 TTGGAACTGTGACAGAAAGCCGCCCAACACTTCAGAAAACTCGTGTCAATCAAAATC 180
Db 121 TTGGAACTGTGACAGAAAGCCGCCCAACACTTCAGAAAACTCGTGTCAATCAAAATC 180
QY 181 CAGATCTGGATGACACCAAGAGGACATTTGAAGTTCCAAAGAGCTCTCGGGAAAGTG 240
Db 181 CAGATCTGGATGACACCAAGAGGACATTTGAAGTTCCAAAGAGCTCTCGGGAAAGTG 240
QY 241 CTGCTGGAGTGTGCAACCACTCAACCTGTGGAAAGTGACATTTTGGCCCTCGAG 300
Db 241 CTGCTGGAGTGTGCAACCACTCAACCTGTGGAAAGTGACATTTTGGCCCTCGAG 300
QY 301 TTTCCTGATCACAAAAAGATCAAGGTGTGGCTGATCTCTTAAACCATTGTGAACAG 360
Db 301 TTTCCTGATCACAAAAAGATCAAGGTGTGGCTGATCTCTTAAACCATTGTGAACAG 360
QY 361 ATTGAAGGCCAAAGCAGTTGTTAAAGTTGTGTGAATTTCTTTCGCTGACAC 420
Db 361 ATTGAAGGCCAAAGCAGTTGTTAAAGTTGTGTGAATTTCTTTCGCTGACAC 420
QY 421 ACACAACTCCAAAGAAACACCAAGTACCTGTGCGCTGAGGTGAAGACAGACTTG 480
Db 421 ACACAACTCCAAAGAAACACCAAGTACCTGTGCGCTGAGGTGAAGACAGACTTG 480
QY 481 GCTCAAGGAGGTGTGACGTGTAATGACACAGCGCAGCTCTCTTGAATTCACATTTG 540
Db 481 GCTCAAGGAGGTGTGACGTGTAATGACACAGCGCAGCTCTCTTGAATTCACATTTG 540
QY 541 CAATCTGAGATTGGGATTTTGTATGAAGCTTGGACAGAGACACTTACGAAAAATTA 600
Db 541 CAATCTGAGATTGGGATTTTGTATGAAGCTTGGACAGAGACACTTACGAAAAATTA 600
QY 601 TACATCTCTCAGAAAGACGACCTAAGAGACAAATGTGGAAATTTACATACCAATC 660
Db 601 TACATCTCTCAGAAAGACGACCTAAGAGACAAATGTGGAAATTTACATACCAATC 660
QY 661 GGAACAACACAGAGAAATCAGATTTCAAGCTCTTGAAGTTGCCCGCTGAGATG 720
Db 661 GGAACAACACAGAGAAATCAGATTTCAAGCTCTTGAAGTTGCCCGCTGAGATG 720
QY 721 TATGAATCGGTGACACCGGACAGAGACAGGAAAGGACGAAATCAATCTGCCGTT 780
Db 721 TATGAATCGGTGACACCGGACAGAGACAGGAAAGGACGAAATCAATCTGCCGTT 780
QY 781 GCCAACAAGGAAATCTAAGTTTCAAGGTTTCAATGAATCAATGCCCTTCAATGAGC 840
Db 781 GCCAACAAGGAAATCTAAGTTTCAAGGTTTCAATGAATCAATGCCCTTCAATGAGC 840
QY 841 AAGTGCAGAGTGTGAGTTCAAGAGAAAGCGCTTCTCATCAAGCTCCGCGCAGATGC 900
Db 841 AAGTGCAGAGTGTGAGTTCAAGAGAAAGCGCTTCTCATCAAGCTCCGCGCAGATGC 900
QY 901 AATAGTGCATCAAGATACCTTGGAAATTCATGATGCGCAATCTCGGAAATTCGCAAGTCC 960
Db 901 AATAGTGCATCAAGATACCTTGGAAATTCATGATGCGCAATCTCGGAAATTCGCAAGTCC 960
QY 961 TTCTGAAAATCTGTGTGAACATCAATGCCCTTCTTGAATTTTGAAGGCCCAACCA 1020
Db 961 TTCTGAAAATCTGTGTGAACATCAATGCCCTTCTTGAATTTTGAAGGCCCAACCA 1020
QY 1021 AAGCCCAAGCCGCTCTTTAGCCGAGGCTCATATTCGTTCAAGTGTGCACTCAG 1080
Db 1021 AAGCCCAAGCCGCTCTTTAGCCGAGGCTCATATTCGTTCAAGTGTGCACTCAG 1080
QY 1081 AAGCAGTTCTGACATATGTTAAAGAGAGACATTAAGAGTGCAGTTGAAGGAAG 1140
Db 1081 AAGCAGTTCTGACATATGTTAAAGAGAGACATTAAGAGTGCAGTTGAAGGAAG 1140

Db 1081 AAGCAGTTCTGACATATGTTAAAGAGAGACATTAAGAGTGCAGTTGAAGGAAG 1140
QY 1141 CACAGCAAGATTCAATCTTATCCGAGGCTTGCTTCAAGCCTTACAGAACTGAATTCGGA 1200
Db 1141 CACAGCAAGATTCAATCTTATCCGAGGCTTGCTTCAAGCCTTACAGAACTGAATTCGGA 1200
QY 1201 GTGCTGAGCAGTCTCAGCAGAGCACAGCCTTACATTTGGAAAGGTGCCGAATCTCCA 1260
Db 1201 GTGCTGAGCAGTCTCAGCAGAGCACAGCCTTACATTTGGAAAGGTGCCGAATCTCCA 1260
QY 1261 GGGGGCCAGAGTGTCCGCGCAGAGAAAGAAACCGAAGTTTCCGCGGGGAGGCCGGGCTCG 1320
Db 1261 GGGGGCCAGAGTGTCCGCGCAGAGAAAGAAACCGAAGTTTCCGCGGGGAGGCCGGGCTCG 1320
QY 1321 CACCCGAGCCCTGCGCCGAGAGAAAGCCCGCGGGTTAACAGCAGGCGGACCGACCC 1380
Db 1321 CACCCGAGCCCTGCGCCGAGAGAAAGCCCGCGGGTTAACAGCAGGCGGACCGACCC 1380
QY 1381 TCGGCGCCACAGAGAAAGAGAGAGGTCTTTAAGATAGGACCCAGCAGATTAACCT 1440
Db 1381 TCGGCGCCACAGAGAAAGAGAGAGGTCTTTAAGATAGGACCCAGCAGATTAACCT 1440
QY 1441 CAGCCCCCGCAGCAGACAGAGCTCTCTGACTGCACTCTTCACTTTCCGAGCTGCT 1500
Db 1441 CAGCCCCCGCAGCAGACAGAGCTCTCTGACTGCACTCTTCACTTTCCGAGCTGCT 1500
QY 1501 GTGAATCTGCAAGGAGGAGTGGCCCTGCAACGTGACCTTGTCTCCCAACTGAGCCCC 1560
Db 1501 GTGAATCTGCAAGGAGGAGTGGCCCTGCAACGTGACCTTGTCTCCCAACTGAGCCCC 1560
QY 1561 GACACCAAGCAGGCTCTCTCTTATCAGCCGCTGTGAATGACCAAGCCTGCCCGCG 1620
Db 1561 GACACCAAGCAGGCTCTCTCTTATCAGCCGCTGTGAATGACCAAGCCTGCCCGCG 1620
QY 1621 ACGACGATGAGAGTGAAGGCGCGAGAGAGATTTCCCACTGAATTAAGCGATTCAT 1680
Db 1621 ACGACGATGAGAGTGAAGGCGCGAGAGAGATTTCCCACTGAATTAAGCGATTCAT 1680
QY 1681 GCTAAGAAAGTGTCTACCAACGAGCGAAATATCTGAAGATCTCGAAGTTTCACTTCG 1740
Db 1681 GCTAAGAAAGTGTCTACCAACGAGCGAAATATCTGAAGATCTCGAAGTTTCACTTCG 1740
QY 1741 TGGTTTCAAGACACATGACCAAGAGACCGCATGCCGAAAGCACTGAAAGTCTCAT 1800
Db 1741 TGGTTTCAAGACACATGACCAAGAGACCGCATGCCGAAAGCACTGAAAGTCTCAT 1800
QY 1801 TTCCCGAATTTTGAACCTTTGACAAATTTTCACTAATTTTCTCAAGAAATTAAGCA 1860
Db 1801 TTCCCGAATTTTGAACCTTTGACAAATTTTCACTAATTTTCTCAAGAAATTAAGCA 1860
QY 1861 CGACTTGCCTGTGGGAAAGCCGCTCAATGCGCAAAATCAGAGATTAACCAAGAAATCGGC 1920
Db 1861 CGACTTGCCTGTGGGAAAGCCGCTCAATGCGCAAAATCAGAGATTAACCAAGAAATCGGC 1920
QY 1921 GATGTATGTCTGAAGAAATTCAGGGCATGAAGCACTGCGCGCTCACTGTGGAAGCAC 1980
Db 1921 GATGTATGTCTGAAGAAATTCAGGGCATGAAGCACTGCGCGCTCACTGTGGAAGCAC 1980
QY 1981 AGCAGAGCCTTGAAGGCGCTGAGAAATGAATCAAGAGCTCCCGGCGGTGGAAGAACTTC 2040
Db 1981 AGCAGAGCCTTGAAGGCGCTGAGAAATGAATCAAGAGCTCCCGGCGGTGGAAGAACTTC 2040
QY 2041 TGCAGAGACTTGAAGTGTGTTAATCAACCGCTCAACACTTCTCTCTGCGG 2100
Db 2041 TGCAGAGACTTGAAGTGTGTTAATCAACCGCTCAACACTTCTCTCTGCGG 2100
QY 2101 CCACTGACACCGGCTCATGACCTACAGAGGTCTGTGAGCGGCTGTGCAACACACCCG 2160
Db 2101 CCACTGACACCGGCTCATGACCTACAGAGGTCTGTGAGCGGCTGTGCAACACACCCG 2160
QY 2161 CCGAGCCACGCGCACTTCAAGGAGCTGCGAGCGCTTTGGCAGAGATCACGAGATGGTG 2220
Db 2161 CCGAGCCACGCGCACTTCAAGGAGCTGCGAGCGCTTTGGCAGAGATCACGAGATGGTG 2220

QY 2221 GCACAGCTCCACGGGTACGATGATCAAGTGGAAATTTCCAGAAAGCTGCAAGAACTCAAG 2280
 DB 2221 GCACAGCTCCACGGGTACGATGATGAGAAATTTCCAGAAAGCTGCAAGAACTCAAG 2280
 QY 2281 AAGATTTGATTTGGCATTTGCAATCTTGTGTTCCGGGAAAGGAGTTCACTCCCTGAGG 2340
 DB 2281 AAGATTTGATTTGGCATTTGCAATCTTGTGTTCCGGGAAAGGAGTTCACTCCCTGAGG 2340
 QY 2341 AGCCTGACGAAAGCTCTGGGGAAAGGGGCTCAGAGAGCATGTTCTTCTGTTCAAGAC 2400
 DB 2341 AGCCTGACGAAAGCTCTGGGGAAAGGGGCTCAGAGAGCATGTTCTTCTGTTCAAGAC 2400
 QY 2401 GTCTCTGATACGAGAGCGGGGGCTGAGAGGCTCCAAATAGTTTAAAGTCCAGGGGAG 2460
 DB 2401 GTCTCTGATACGAGAGCGGGGGCTGAGAGGCTCCAAATAGTTTAAAGTCCAGGGGAG 2460
 QY 2461 CTCCCGCTCTATGGCATGACGATTTGAGAGAGCGAAAGCATGAGGGGGTCCCACTGC 2520
 DB 2461 CTCCCGCTCTATGGCATGACGATTTGAGAGAGCGAAAGCATGAGGGGGTCCCACTGC 2520
 QY 2521 CTGACCTCTCGGGGGCAGCGGCGATCCATCTGTGGCGGCACTTCTGGTCCGAGATG 2580
 DB 2521 CTGACCTCTCGGGGGCAGCGGCGATCCATCTGTGGCGGCACTTCTGGTCCGAGATG 2580
 QY 2581 GAGAACTGGGTGAGAGCATCCAGATGGCCATTGACCTGGCGGAGAAAGAGACAGACCCC 2640
 DB 2581 GAGAACTGGGTGAGAGCATCCAGATGGCCATTGACCTGGCGGAGAAAGAGACAGACCCC 2640
 QY 2641 GCCCTGAGTCTCTGGCCAGCAGCCCCCTGACAAAGTCCCCCTGATGAAAGCAGCCG 2700
 DB 2641 GCCCTGAGTCTCTGGCCAGCAGCCCCCTGACAAAGTCCCCCTGATGAAAGCAGCCG 2700
 QY 2701 GCTGACCAAGAGTCAAGAGATGACTGAGCGCTTCGGCAGCATGCTGTGAGGCCAGGCC 2760
 DB 2701 GCTGACCAAGAGTCAAGAGATGACTGAGCGCTTCGGCAGCATGCTGTGAGGCCAGGCC 2760
 QY 2761 CGGACCGGGGGCAACAATGTCAGTGTGTCGACCGGCAACCGAGCGTCCATG 2820
 DB 2761 CGGACCGGGGGCAACAATGTCAGTGTGTCGACCGGCAACCGAGCGTCCATG 2820
 QY 2821 GTGGAATTCAGCATCCGAGTGAAGAAATCACTTGTCTGGAACCTGCTGAGGAATTCAAA 2880
 DB 2821 GTGGAATTCAGCATCCGAGTGAAGAAATCACTTGTCTGGAACCTGCTGAGGAATTCAAA 2880
 QY 2881 AACAGCAAGGGTGGCAAGGCTGTGGGTGTTCACAACTTCTGCTGTTCTTCTAC 2940
 DB 2881 AACAGCAAGGGTGGCAAGGCTGTGGGTGTTCACAACTTCTGCTGTTCTTCTAC 2940
 QY 2941 AATTCACACAGGACAAATCACTCCCTTGCAGGCTGCTGCTCGGCTACCTGCTCAC 3000
 DB 2941 AATTCACACAGGACAAATCACTCCCTTGCAGGCTGCTGCTCGGCTACCTGCTCAC 3000
 QY 3001 ATCCCTCTGAGTCCGAGAAATCACTAGTGTTCAGAGTGCACCTTCAAGTCC 3060
 DB 3001 ATCCCTCTGAGTCCGAGAAATCACTAGTGTTCAGAGTGCACCTTCAAGTCC 3060
 QY 3061 CACGTCTACTACTTCAAGGGCGGAAAGGAGTACAGTTTCGAAAGGTGATGAAAGTATC 3120
 DB 3061 CACGTCTACTACTTCAAGGGCGGAAAGGAGTACAGTTTCGAAAGGTGATGAAAGTATC 3120
 QY 3121 CGAGAGGCAACAGCTCTGCTGCGAGCCCAAGTGTGAGCAACAAAGGTCTCTGAG 3180
 DB 3121 CGAGAGGCAACAGCTCTGCTGCGAGCCCAAGTGTGAGCAACAAAGGTCTCTGAG 3180
 QY 3181 TATTGATGGCCGAGACACTCTGTTCCGAGTGGCTTCTTCTGGAAGAGCTTCTCTT 3240
 DB 3181 TATTGATGGCCGAGACACTCTGTTCCGAGTGGCTTCTTCTGGAAGAGCTTCTCTT 3240
 QY 3241 CTCTGTATTATGAAGCTGTGTTAAATTAACCTGTCTGAAATCAAAAATCATGGCTT 3300
 DB 3241 CTCTGTATTATGAAGCTGTGTTAAATTAACCTGTCTGAAATCAAAAATCATGGCTT 3300

QY 3301 CCAGAGGCTCTCTGTCCTCACAAGCGCGGTTTTTAAACCCGAGCTCTCAGCGTTTGA 3360
 DB 3301 CCAGAGGCTCTCTGTCCTCACAAGCGCGGTTTTTAAACCCGAGCTCTCAGCGTTTGA 3360
 QY 3361 TGAACAGCGCTCCACCTTCAGTCTGAGCATCGCTGGGGGGCTGTTCTTTAGCTAGTG 3420
 DB 3361 TGAACAGCGCTCCACCTTCAGTCTGAGCATCGCTGGGGGGCTGTTCTTTAGCTAGTG 3420
 QY 3421 CCAGTATTAAACATTGTCAAT 3442
 DB 3421 CCAGTATTAAACATTGTCAAT 3442
 RESULT 3
 CS130723
 LOCUS CS130723 5028 bp DNA linear PAT 02-AUG-2005
 DEFINITION Sequence 9 from Patent WO2005064009.
 ACCESSION CS130723
 VERSION CS130723.1 GI:71792793
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homidae; Homo.
 REFERENCE 1
 AUTHORS Oernboft, T. F.
 TITLES Classification of cancer
 JOURNAL Patent: WO 2005064009-A 9 14-JUL-2005;
 Aros Applied Biotechnology APS (DK)
 FEATURES
 source 1..5028
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 /note="FERM, RhoGEF (ARHGEF) and pleckstrin domain pro"
 ORIGIN
 Query Match 99.9%; Score 3438.8; DB 2; Length 5028;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 3440; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 CGCCGAGCGCGCGCGCTGTGAGATATCTTAAGCGCTTTCATGAGGAGAAAT 60
 DB 289 CGCCGAGCGCGCGCGCTGTGAGATATCTTAAGCGCTTTCATGAGGAGAAAT 348
 QY 61 GAGCAGAGGCGGACCCAGAGATCAAGCTGGGGGCCCGGAAATTCGGGGATCACTACC 120
 DB 349 GAGCAGAGGCGGACCCAGAGATCAAGCTGGGGGCCCGGAAATTCGGGGATCACTACC 408
 QY 121 TTGGAACGTGACAGAAAGCGGCCCAACCTTCAGGAAATCTGTGTCATCAAAATC 180
 DB 409 TTGGAACGTGACAGAAAGCGGCCCAACCTTCAGGAAATCTGTGTCATCAAAATC 468
 QY 181 CAGATCTGATGACACCCAGAGAGGATTTGAAGTTCCACAAAGAGCTCTGAGGAGGTG 240
 DB 469 CAGATCTGATGACACCCAGAGAGGATTTGAAGTTCCACAAAGAGCTCTGAGGAGGTG 528
 QY 241 CTGCTGATGACATTTGCAACCACTCACTGTGGAAGGTGACTATTTGGCTCGAG 300
 DB 529 CTGCTGATGACATTTGCAACCACTCACTGTGGAAGGTGACTATTTGGCTCGAG 588
 QY 301 TTTCTGATCAAAAAAGATCAAGGTGAGTCTCTTAAACCAATGGAACAG 360
 DB 589 TTTCTGATCAAAAAAGATCAAGGTGAGTCTCTTAAACCAATGGAACAG 648
 QY 361 ATTAGAAGCCAAAGCACTGTTGTTAAGTTGTGTAATTTCTTCCGCTGACAC 420
 DB 649 ATTAGAAGCCAAAGCACTGTTGTTAAGTTGTGTAATTTCTTCCGCTGACAC 708
 QY 421 ACACAACTCCAAAGAACTCAACAGATACCTGTTGCGCTGAGGTGAAGCAAGACTTG 480
 DB 709 ACACAACTCCAAAGAACTCAACAGATACCTGTTGCGCTGAGGTGAAGCAAGACTTG 768

481 GCTCAAGGAGGTTGACGTGTAATGACACAGCGAGCTCTCTGATTTTACAGATTG 540
| | | | |
769 GCTCAAGGAGGTTGACGTGTAATGACACAGCGAGCTCTCTGATTTTACAGATTG 828
| | | | |
541 CAATCGAGATTGGGATTTTGTATGAAAGCTTGAACAGAGCATTTAGCAAAATAA 600
| | | | |
829 CAATCGAGATTGGGATTTTGTATGAAAGCTTGAACAGAGCATTTAGCAAAATAA 888
| | | | |
601 TACATTCCTCAGACAGACGCTAGAGGACAAATGTGGAATTTTACATAACCAT 660
| | | | |
889 TACATTCCTCAGACAGACGCTAGAGGACAAATGTGGAATTTTACATAACCAT 948
| | | | |
661 GGAACAAACACAGAGAAATCAGATTTCAGGCTCTAGAGATTGCCGCGGCTAGAGAT 720
| | | | |
949 GGAACAAACACAGAGAAATCAGATTTCAGGCTCTAGAGATTGCCGCGGCTAGAGAT 1008
| | | | |
721 TATGGAATCCGGTTGACCCGGCCAAAGACAGGGAAGGACGAAATCAATCTGGCGTT 780
| | | | |
1009 TATGGAATCCGGTTGACCCGGCCAAAGACAGGGAAGGACGAAATCAATCTGGCGTT 1068
| | | | |
781 GCCAACAACGGAATTTCTAGTGTTCAGGGTTTCACTAAGATCAATGCTTCAACTGGG 840
| | | | |
1069 GCCAACAACGGAATTTCTAGTGTTCAGGGTTTCACTAAGATCAATGCTTCAACTGGG 1128
| | | | |
841 AAGGTCCGGAAGCTGAGCTTCAAGAGAAAGCGCTTCTCATCAAGCTCCGGCCAGATG 900
| | | | |
1129 AAGGTCCGGAAGCTGAGCTTCAAGAGAAAGCGCTTCTCATCAAGCTCCGGCCAGATG 1188
| | | | |
901 AATAGTCCGTACAGAGATACCTTGAATTCCTGATGGCAGATCGGGATTTCTGCAAGT 960
| | | | |
1189 AATAGTCCGTACAGAGATACCTTGAATTCCTGATGGCAGATCGGGATTTCTGCAAGT 1248
| | | | |
961 TTCTGGAATAATCTGTGTTGAACATCATGCTTCTTGAACCTTTTGAAGGCCCAACCA 1020
| | | | |
1249 TTCTGGAATAATCTGTGTTGAACATCATGCTTCTTGAACCTTTTGAAGGCCCAACCA 1308
| | | | |
1021 AAGCCCAAGCCGCTCTTTAGCCGGGGATCATCTTCCGTTCAAGTGGTGGATCTGAG 1080
| | | | |
1309 AAGCCCAAGCCGCTCTTTAGCCGGGGATCATCTTCCGTTCAAGTGGTGGATCTGAG 1368
| | | | |
1081 AAGCAGGTTCTGACTATGTATTAAGAGAGACATTAAGAGGTGATGTTTGAAGAGAG 1140
| | | | |
1369 AAGCAGGTTCTGACTATGTATTAAGAGAGACATTAAGAGGTGATGTTTGAAGAGAG 1428
| | | | |
1141 CACAGTAATATTCATTCTATCCGAGACCTTGTCTTCAACCTTACAGAACTGAATTCGGA 1200
| | | | |
1429 CACAGTAATATTCATTCTATCCGAGACCTTGTCTTCAACCTTACAGAACTGAATTCGGA 1488
| | | | |
1201 GTGCTGAGACAGTCTCAGACAGACACAGCTTACATTTGAGAGAGTCCGAAATCTCA 1260
| | | | |
1489 GTGCTGAGACAGTCTCAGACAGACACAGCTTACATTTGAGAGAGTCCGAAATCTCA 1548
| | | | |
1261 GGGGGGACAGAGCTGCGGCGAGAGAAAGAACCGAAGTTCGCGCGGGAGCGCGGGTCTG 1320
| | | | |
1549 GGGGGGACAGAGCTGCGGCGAGAGAAAGAACCGAAGTTCGCGCGGGAGCGCGGGTCTG 1608
| | | | |
1321 CACCCGAGGCTGCGCGGAGAGAAAGCCCGCGGGTAAACAAGAGCGGGAACGAGCGG 1380
| | | | |
1609 CACCCGAGGCTGCGCGGAGAGAAAGCCCGCGGGTAAACAAGAGCGGGAACGAGCGG 1668
| | | | |
1381 TCGGCGCCACAGAGAGAGAGAGAGTCTTAAAGATAGAACCCAGACAGTAAACT 1440
| | | | |
1669 TCGGCGCCACAGAGAGAGAGAGAGTCTTAAAGATAGAACCCAGACAGTAAACT 1728
| | | | |
1441 CAGCCCGGACAGCAAGACAGAGCTCTGACTGAGAGTCTTCACTTTCCAGAGCTGTCT 1500
| | | | |
1729 CAGCCCGGACAGCAAGACAGAGCTCTGACTGAGAGTCTTCACTTTCCAGAGCTGTCT 1788
| | | | |
1501 GTGAATCTCCAGAGGAGAGATGGCCCTTGCAACGTGACCTTGTCTCCCAACTGAGCC 1560
| | | | |
1789 GTGAATCTCCAGAGGAGAGATGGCCCTTGCAACGTGACCTTGTCTCCCAACTGAGCC 1848
| | | | |

1561 GACACCAAGACAGGCTCTCTCTTGTATCAGCCGCTGTGAATGACAGAGCTGCCCGG 1620
| | | | |
1849 GACACCAAGACAGGCTCTCTCTTGTATCAGCCGCTGTGAATGACAGAGCTGCCCGG 1908
| | | | |
1621 ACGAGCATGAGATGAGGCGCGGAGAGAGATTTCCCATGATTAAGCTTCTATTA 1680
| | | | |
1909 ACGAGCATGAGATGAGGCGCGGAGAGAGATTTCCCATGATTAAGCTTCTATTA 1968
| | | | |
1681 GCTAAGAGAGTCTACACAGGAGGACATATCTGAAGAGATCTCGAAGTTATCATCT 1740
| | | | |
1969 GCTAAGAGAGTCTACACAGGAGGACATATCTGAAGAGATCTCGAAGTTATCATCT 2028
| | | | |
1741 TGGTTTCAGACACAGTAGAGCAAGAGAGACGCAATGCCGGAAGCATGAAAAGTCTCAT 1800
| | | | |
2029 TGGTTTCAGACACAGTAGAGCAAGAGAGACGCAATGCCGGAAGCATGAAAAGTCTCAT 2088
| | | | |
1801 TTCCGGAATTTTGAACCTTTGCAAAATTTTCACTAATTTTCTCAGAGAAATGAGCA 1860
| | | | |
2089 TTCCGGAATTTTGAACCTTTGCAAAATTTTCACTAATTTTCTCAGAGAAATGAGCA 2148
| | | | |
1861 CGACTTGGCCCTGTGGGAAGGCGGCTCAAAATGCCAAATCAGAGATTACAAAGATCGG 1920
| | | | |
2149 CGACTTGGCCCTGTGGGAAGGCGGCTCAAAATGCCAAATCAGAGATTACAAAGATCGG 2208
| | | | |
1921 GATGTCATGCTGAAGAACATTCAGGSCATGAAGCATCTGGCGCTCACCTGTGAAGAC 1980
| | | | |
2209 GATGTCATGCTGAAGAACATTCAGGSCATGAAGCATCTGGCGCTCACCTGTGAAGAC 2268
| | | | |
1981 AGCGAGGCTTTGAGAGGCGCTTGAAGATGGAATCAAGAGCTCCGGCGCTGGAACAT 2040
| | | | |
2269 AGCGAGGCTTTGAGAGGCGCTTGAAGATGGAATCAAGAGCTCCGGCGCTGGAACAT 2328
| | | | |
2041 TGCAGAGCTTTGAGCTCAGAGAGGTGTTTACCTACAGGCTCAACCTTCTCTGCGG 2100
| | | | |
2329 TGCAGAGCTTTGAGCTCAGAGAGGTGTTTACCTACAGGCTCAACCTTCTCTGCGG 2388
| | | | |
2101 CCACTGACACCGGCTCATGACCTACAGAGGCTCTGAGACGGCTGTGCAAAACCAACCG 2160
| | | | |
2389 CCACTGACACCGGCTCATGACCTACAGAGGCTCTGAGACGGCTGTGCAAAACCAACCG 2448
| | | | |
2161 CCGAGCCACGCGGACTTCAAGGATCTGCCAGACCGCTTTGGCAGAGATCAACGAGATG 2220
| | | | |
2449 CCGAGCCACGCGGACTTCAAGGATCTGCCAGACCGCTTTGGCAGAGATCAACGAGATG 2508
| | | | |
2221 GCACAGCTCCACGATGATGATCAAGATGAGAAATTTCCAGAGCTGACAGAACTCAAG 2280
| | | | |
2509 GCACAGCTCCACGATGATGATCAAGATGAGAAATTTCCAGAGCTGACAGAACTCAAG 2568
| | | | |
2281 AAAGATTTGATGATGATGATCAATCTTGTGTTCCGGGAAGGAGTTCATCGTCTGG 2340
| | | | |
2569 AAAGATTTGATGATGATGATCAATCTTGTGTTCCGGGAAGGAGTTCATCGTCTGG 2628
| | | | |
2341 AGCTTCAGCAAGCTTCTGGGGAGAGGGGCTCCAGACAGGCAATGTTTCTCTGTTCAAC 2400
| | | | |
2629 AGCTTCAGCAAGCTTCTGGGGAGAGGGGCTCCAGACAGGCAATGTTTCTCTGTTCAAC 2688
| | | | |
2401 GTCTGCTATATCAACAGAGCGGGGCTGACGGCTCTCAATTCAGTTTAAAGTCCAGGG 2460
| | | | |
2689 GTCTGCTATATCAACAGAGCGGGGCTGACGGCTCTCAATTCAGTTTAAAGTCCAGGG 2748
| | | | |
2461 CTCCGCTCTATATGATGATGATGATGAGAGAGAGAGAGAGATGAGGGGTGCCCATCTG 2520
| | | | |
2749 CTCCGCTCTATATGATGATGATGATGAGAGAGAGAGAGATGAGGGGTGCCCATCTG 2808
| | | | |
2521 CTGACCTCTCCGGGACAGCGGCAATCTATCTGTGACCGGCAATCTCTGCTCCGAGATG 2580
| | | | |
2809 CTGACCTCTCCGGGACAGCGGCAATCTATCTGTGACCGGCAATCTCTGCTCCGAGATG 2868
| | | | |
2581 GAGAGTGGGTTGAGAGCATTCAGATGAGCTTGAACCTGGGAGAGAGAGAGAGAGAG 2640
| | | | |
2869 GAGAGTGGGTTGAGAGCATTCAGATGAGCTTGAACCTGGGAGAGAGAGAGAGAGAG 2928
| | | | |
2641 GCCCTGAGTTCTGGCCAGAGAGCCCTGTAACAAGATCCCTGATGAAGCCACCGG 2700
| | | | |

Db	2929	GCCTCTAGTTCCTGCGCCAGACGCCCCCTGACAAAGTCCCTGATGAAGACACCGCC	2988
Qy	2701	GCTGACCAAGAGTCAAGAGATGACCTGAGCGCCTTCGCGCACATCGCTGAGCGCCAGGCC	2766
Db	2989	GCTGACCAAGAGTCAAGAGATGACCTGAGCGCCTTCGCGCACATCGCTGAGCGCCAGGCC	3048
Qy	2761	CCGCAACCGCGGCACACAAATGTCGACGCTGTGCTGGCACCCGCAACCAAGGCTCTCCATG	2820
Db	3049	CCGCAACCGCGGCACACAAATGTCGACGCTGTGCTGGCACCCGCAACCAAGGCTCTCCATG	3108
Qy	2821	GTGGAATTCAGCATCGCAGTGGAGAAATGATGTGTCTGGAACCTGCTGAGGAATTCAA	2888
Db	3109	GTGGAATTCAGCATCGCAGTGGAGAAATGATGTGTCTGGAACCTGCTGAGGAATTCAA	3168
Qy	2881	AACAGCAACGGGTGGCAGAAAGCTGTGGTGTGTTCAAACTTCTGCTGTCTTCTAC	2940
Db	3169	AACAGCAACGGGTGGCAGAAAGCTGTGGTGTGTTCAAACTTCTGCTGTCTTCTAC	3228
Qy	2941	AAATCACACCAAGCAATCATCCCTTTCGCGACGCTGCTCTGCTGGTACTTCGCTCAC	3000
Db	3229	AAATCACACCAAGCAATCATCCCTTTCGCGACGCTGCTCTGCTGGTACTTCGCTCAC	3288
Qy	3001	ATCCCTCTGAGTCCGAGAAATCATCCCTTTCGCGACGCTGCTCTGCTGGTACTTCGCTCAC	3066
Db	3289	ATCCCTCTGAGTCCGAGAAATCATCCCTTTCGCGACGCTGCTCTGCTGGTACTTCGCTCAC	3348
Qy	3061	CACGCTTACTACTTCAAGGGCGGAAAGCAGTACAGTTTCAAAAGTGATGAAGTATC	3120
Db	3349	CACGCTTACTACTTCAAGGGCGGAAAGCAGTACAGTTTCAAAAGTGATGAAGTATC	3408
Qy	3121	CGCAGTGCACACAGCTCTGCTGCGGACCCCACTGTTGAGCCACAAAGATCTCTGTG	3180
Db	3409	CGCAGTGCACACAGCTCTGCTGCGGACCCCACTGTTGAGTCAAAAGATCTCTGTG	3468
Qy	3181	TATGTAAGGCGCGGACACATCGTGTTCGCGAGTGGCTGTTTCTCGGAAGAGTTCTCTT	3240
Db	3469	TATGTAAGGCGCGGACACATCGTGTTCGCGAGTGGCTGTTTCTCGGAAGAGTTCTCTT	3528
Qy	3241	CTTCTGTATTATATAGAGCTGTGTAATAACCTGTCTGGAATCAAAACATAGCTT	3300
Db	3529	CTTCTGTATTATATAGAGCTGTGTAATAACCTGTCTGGAATCAAAACATAGCTT	3588
Qy	3301	CCGACAGCTCTCTGTCTCCACAGCGCGGTTTTTAAACCCGACCTTCAAGGTTTGA	3366
Db	3589	CCGACAGCTCTCTGTCTCCACAGCGCGGTTTTTAAACCCGACCTTCAAGGTTTGA	3648
Qy	3361	TGAACAGCGCTCCCACTCCAGTCTGTGCATCCGCTGGGGGCGCTGTTCTTTACCTAGT	3420
Db	3649	TGAACAGCGCTCCCACTCCAGTCTGTGCATCCGCTGGGGGCGCTGTTCTTTACCTAGT	3708
Qy	3421	CCAGTATTAAACATTTGTCATT	3442
Db	3709	CCAGTATTAAACATTTGTCATT	3730
RESULT 4			
BD231200		4687 bp	DNA
LOCUS		BD231200	Linear
DEFINITION		Human cytoskeleton associated proteins.	PAT 17-JUL-2003
ACCESSION		BD231200.1	GI:33040970
VERSION		JP 2002526076-A/2.	
KEYWORDS		Homo sapiens (human)	
SOURCE		Homo sapiens	
ORGANISM		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.	
REFERENCE		1. (bases 1 to 4687)	
AUTHORS		Lal, P., Tang, T.Y., Yue, H., Hillman, J.L., Bandman, O., Corley, N.C., Guegler, K.J., Patterson, C., Azimtai, Y. and Baughn, M.R.	
TITLE		Human cytoskeleton associated proteins	
JOURNAL		Patent: JP 2002526076-A 2 20-AUG-2002;	

COMMENT	INCYTE PHARMACEUTICALS INC
OS Homo sapiens (human)	
PN JP 2002526076-A/2	
PD 20-AUG-2002	
PP 17-SEP-1999 JP 2000574254	
PR 18-SEP-1998 US 60/172226, 27-APR-1999 US 60/131321	
PREFTI LAL, TOM Y TANG, HENRY YUE, JENNIFER L HILLMAN, OLGA PI BANDMAN,	
PI NEIL C CORLEY, KARL J GUESLER, CHANDRA PATTERSON, YALDA AZIMZAI,	
PI MARIN R BAUGHN	
PC C12N15/09,A61K38/00,A61K45/00,A61P1/16,A61P3/00,A61P3/06 PC	
,A61P3/10,A61P5/00,	
PC A61P5/38,A61P7/00,A61P7/06,A61P9/00,A61P9/10,A61P11/06 PC	
,A61P13/00,A61P3/12,	
PC A61P17/06,A61P19/02,A61P19/10,A61P21/04,A61P25/00,A61P25/00,	
PC A61P25/14,A61P25/16,A61P25/28,A61P29/00,A61P31/04,A61P31/12,	
PC A61P31/18,	
PC A61P33/00,A61P35/00,A61P35/02,A61P37/00,A61P43/00,C07K14/47,	
PC C07K16/18,	
PC C12N1/15,C12N1/19,C12N1/21,C12N5/10,C12P21/02,C12P21/08 PC	
,C12O1/68,G01N33/15,	
PC G01N33/50,G01N33/53,G01N33/53,G01N33/56,G01N33/56,G01N33/56,G01N33/56,	
PC G01N37/02	
CC Inocyte ID No: 1320252	
FT Key	
FT source	
FT location/Qualifiers	
1. .4687	
/organism="Homo sapiens"	
/mol_type="genomic DNA"	
/db_xref="taxon:9606"	
ORIGIN	
Query Match	
Best Local Similarity	
Matches 3439; Conservative	
99.9%; Score 3437.2; DB 2; Length 4687;	
99.9%; Pred. No. 0;	
0; Mismatches 3; Indels 0; Gaps 0	
1 CGCCGACGCGCGCGCTGTGAGATATTTCTTAAGCGGCTTTCATCATGAGAAATA 60	
42 CGCGGACGCGCGCGCTGTGAGATATTTCTTAAGCGGCTTTCATCATGAGAAATA 101	
61 GAGCAGAGGCGGACCCGACGATCAAGCTGGGGGCGCCGGAATAATTCGGGATCAGTACC 120	
102 GAGCAGAGGCGGACCCGACGATCAAGCTGGGGGCGCCGGAATAATTCGGGATCAGTACC 161	
121 TTGGAAGGTGACAGAGCGCGCCCAACACCTTCAGGAAACCTGTGCATCAAAATC 180	
162 TTGGAAGGTGACAGAGCGCGCCCAACACCTTCAGGAAACCTGTGCATCAAAATC 221	
181 CAGATGCTGATGACACCCAGAGGACATTTGAATTCACAAAGAGCTCTGGGAAGGTG 240	
222 CAGATGCTGATGACACCCAGAGGACATTTGAATTCACAAAGAGCTCTGGGAAGGTG 281	
241 CTGCTGATGAGTTTGCAACCACTCAACCTGTGGAAAGGTGACTATTTGGCCTCGAG 300	
282 CTGCTGATGAGTTTGCAACCACTCAACCTGTGGAAAGGTGACTATTTGGCCTCGAG 341	
301 TTTTCTGATCAGAAAAGATACGCGTGTGATCTCTTAAACCATTGTGAAACG 360	
342 TTTTCTGATCAGAAAAGATACGCGTGTGATCTCTTAAACCATTGTGAAACG 401	
361 ATTGAAGGCGCAAGACGTTGTTGAATTTGAGTAATTTCTTCGCGCTGACAC 420	
402 ATTGAAGGCGCAAGACGTTGTTGAATTTGAGTAATTTCTTCGCGCTGACAC 461	
421 ACACACTCCAGAGAACTCACAAAGTACTGTTCGCGCTGACAGGTGAAGCAGACTTG 480	
462 ACACACTCCAGAGAACTCACAAAGTACTGTTCGCGCTGACAGGTGAAGCAGACTTG 521	
481 GCTCAGGCGAGTTGAGGTGTAATGACACGCGGAGTCTCTTGAATTTACACATTTG 540	

Db 522 GCTCAAGCAGGTTGACGTGTAATGACACAGCGCAGCTCTCTTGATTTCAACATTGTC 581
QY 541 CAATCTGAGATTGGGATTTTGTATGTAAGCTTGAGAGAGAGACTTACCAAAAATAA 600
Db 582 CAATCTGAGATTGGGATTTTGTATGTAAGCTTGAGAGAGAGACTTACCAAAAATAA 641
QY 601 TACATTACTCAGACAGACCACTAGAGACAATAATCGTGGAAATTTTCCATNACCAAT 660
Db 642 TACATTACTCAGAGAACCACTAGAGACAATAATCGTGGAAATTTTCCATNACCAAT 701
QY 661 GGAACAAACACAGAGAAATCAATTTTCCAGCTCTTAAAGATTGCCCCGTGAGATG 720
Db 702 GGAACAAACACAGAAATCAGATTTCCAGCTCTTAAAGATTGCCCCGTGAGATG 761
QY 721 TATGGAATCCGGTTGACCCGGCCAGAGACAGGGAAGGACGAAAGTCAATCTGGCCGT 780
Db 762 TATGGAATCCGGTTGACCCGGCCAGAGAGAGGACGAAAGTCAATCTGGCCGT 821
QY 781 GCCAACACGGGAATTTAGTGTTCAGGGTTTCACTAAGATCAATGCTTCAACTGGGCC 840
Db 822 GCCAACACGGGAATTTAGTGTTCAGGGTTTCACTAAGATCAATGCTTCAACTGGGCC 881
QY 841 AAGGTGCGGAAGTGAAGCTTCAAGAGAAAGCGTTTCTCATCAAGCTCCGGCCAGATGCC 900
Db 882 AAGGTGCGGAAGTGAAGCTTCAAGAGAAAGCGTTTCTCATCAAGCTCCGGCCAGATGCC 941
QY 901 AATPAGTGGTACAGAGATCTTGGAAATCTTGAAATGCGGACGCGGATTTCTGCAATGCC 960
Db 942 AATPAGTGGTACAGAGATCTTGGAAATCTTGAAATGCGGACGCGGATTTCTGCAATGCC 1001
QY 961 TTCTGGAAATCTGTGTGTAACATCATGCTTCTTTAGAATTTTGAAGGCCAAACCA 1020
Db 1002 TTCTGGAAATCTGTGTGTAACATCATGCTTCTTTAGAATTTTGAAGGCCAAACCA 1061
QY 1021 AAGCCCAAGCCCGTCTTTTACGCGGGGTATCATTTTCGGTTCAGTGTGGAAATCCAG 1080
Db 1062 AAGCCCAAGCCCGTCTTTTACGCGGGGTATCATTTTCGGTTCAGTGTGGAAATCCAG 1121
QY 1081 AAGCAGTCTGCACTATGTTAAAGAAAGAGACATPAGAAAGTGTGATTTTGAAGGAAG 1140
Db 1122 AAGCAGTCTGCACTATGTTAAAGAAAGAGACATPAGAAAGTGTGATTTTGAAGGAAG 1181
QY 1141 CACAGCAGATTCATTCTATCCGAGACCTTGCTTCAAGCTTACAGAACTGAATTCGAA 1200
Db 1182 CACAGCAGATTCATTCTATCCGAGACCTTGCTTCAAGCTTACAGAACTGAATTCGAA 1241
QY 1201 GTGCTGAGACAGTCTTACAGAGACACAGCTTACATTTGAGAAAGTGTCCGAATCTCA 1260
Db 1242 GTGCTGAGACAGTCTTACAGAGACACAGCTTACATTTGAGAAAGTGTCCGAATCTCA 1301
QY 1261 GGGGGGCAAGGCTGCGGCGAGAAAGAACCGAAAGTTTCCGCGGGGAGCCGGGGTGC 1320
Db 1302 GGGGGGCAAGGCTGCGGCGAGAAAGAACCGAAAGTTTCCGCGGGGAGCCGGGGTGC 1361
QY 1321 CACCCGAGGCTTGCGCGAGAGAAAGCCCGCGGGTAAACAGACGAGCGAGCGGCC 1380
Db 1362 CACCCGAGGCTTGCGCGAGAGAAAGCCCGCGGGTAAACAGACGAGCGAGCGGCC 1421
QY 1381 TGGGCGGCCCAAGAGAAAGAGAGAGTCTTTAAAGATAGAACCCAGACAGATTAACCT 1440
Db 1422 TGGGCGGCCCAAGAGAAAGAGAGAGTCTTTAAAGATAGAACCCAGACAGATTAACCT 1481
QY 1441 CAGCCCCCGCAGCAAGACAGAGCTCCCTGACCTGGACGTCTCACTTTCCGAGCTGTCT 1500
Db 1482 CAGCCCCCGCAGCAAGACAGAGCTCCCTGACCTGGACGTCTCACTTTCCGAGCTGTCT 1541
QY 1501 GTGAATCTGCAAGGGGAGATGGCCCTTGCAACGTGACCTTGTCTTCCAACTGAGCCCC 1560
Db 1542 GTGAATCTGCAAGGGGAGATGGCCCTTGCAACGTGACCTTGTCTTCCAACTGAGCCCC 1601
QY 1561 GACACCAAGACAGGCTCTCTCTTGAATACGCGGCTGTGAATGACCAAGGCTGCCCCGG 1620
Db 1602 GACACCAAGACAGGCTCTCTCTTGAATACGCGGCTGTGAATGACCAAGGCTGCCCCGG 1661

QY 1621 ACGAGCATGAGATGAGGGCCGAGAGAGAGATTTCCCACTGATTAAGCGTACTTCATA 1680
Db 1662 ACGAGCATGAGATGAGGGCCGAGAGAGAGATTTCCCACTGATTAAGCGTACTTCATA 1721
QY 1681 GCTAAGAAAGTGTCTAACCCAGGCGAAATATCTGAAGAGATCTCGAAGTTATCACTTCG 1740
Db 1722 GCTAAGAAAGTGTCTAACCCAGGCGAAATATCTGAAGAGATCTCGAAGTTATCACTTCG 1781
QY 1741 TGGTTTCAGAGACAGTGAAGCAAAAGAGACCGCATGCGGAAAGCACTGAAAAGTCTCAT 1800
Db 1782 TGGTTTCAGAGACAGTGAAGCAAAAGAGACCGCATGCGGAAAGCACTGAAAAGTCTCAT 1841
QY 1801 TTCCGGAATTTTGAACCTTTGACAAATTTCTAATTTTCTCAAGGAAATGAGCAA 1860
Db 1842 TTCCGGAATTTTGAACCTTTGACAAATTTCTAATTTTCTCAAGGAAATGAGCAA 1901
QY 1861 CGACTTGCCTGTGGAAAGCCGCTCAATGCCCCAATCAGAGATTTCCAAAGAAATCGGC 1920
Db 1902 CGACTTGCCTGTGGAAAGCCGCTCAATGCCCCAATCAGAGATTTCCAAAGAAATCGGC 1961
QY 1921 GATGTCAATGCTGAAGAACTTCAAGGCAATGAGCACCTTGCGGCTCACTGTGAAGCAC 1980
Db 1962 GATGTCAATGCTGAAGAACTTCAAGGCAATGAGCACCTTGCGGCTCACTGTGAAGCAC 2021
QY 1981 AGCAGAGCCTTGAGAGCCCTGGAAGATGGAATCAAGAGCTCCGGCGGCTGGAAGAACTTC 2040
Db 2022 AGCAGAGCCTTGAGAGCCCTGGAAGATGGAATCAAGAGCTCCGGCGGCTGGAAGAACTTC 2081
QY 2041 TGCAGAGCTTTGAGCTGCAAGAGGTGTGTATCTACCGGCTCAACCTTCTCTGCGG 2100
Db 2082 TGCAGAGCTTTGAGCTGCAAGAGGTGTGTATCTACCGGCTCAACCTTCTCTGCGG 2141
QY 2101 CCAGTCACCCGCTCATATGCACTCAAGAGAGCTCTGAGCGGCTGGAACCAACCAACCGG 2160
Db 2142 CCAGTCACCCGCTCATATGCACTCAAGAGAGCTCTGAGCGGCTGGAACCAACCAACCGG 2201
QY 2161 CCGAGCCACGCGCACTTGAGGAGCTGCGAGACCGCTTTTGGCAGAGATCAGAGAGATGCTG 2220
Db 2202 CCGAGCCACGCGCACTTGAGGAGCTGCGAGACCGCTTTTGGCAGAGATCAGAGAGATGCTG 2261
QY 2221 GCACAGCTTCACGCTACATGATCAAGATGAGAAATTTCCAGAGCTGACAGACTCAAG 2280
Db 2262 GCACAGCTTCACGCTACATGATCAAGATGAGAAATTTCCAGAGCTGACAGACTCAAG 2321
QY 2281 AAAGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTG 2340
Db 2322 AAAGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTG 2381
QY 2341 AGCCTCAGCAAGCTCTCGGGGAAAGGGGCTCCAGCAGCGCATGTTCTCTGTTCAACGAC 2400
Db 2382 AGCCTCAGCAAGCTCTCGGGGAAAGGGGCTCCAGCAGCGCATGTTCTCTGTTCAACGAC 2441
QY 2401 GTCTGTATACAGAGCCGGGGGCTTACCGGCTTCAATCAATTTTAAAGTCCACGGGCAAG 2460
Db 2442 GTCTGTATACAGAGCCGGGGGCTTACCGGCTTCAATCAATTTTAAAGTCCACGGGCAAG 2501
QY 2461 CTCCGCTCTATGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2520
Db 2502 CTCCGCTCTATGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2561
QY 2521 CTGACCTTCGGGGGCAAGCGGAGTCAATCATCTGTGGCCGCAAGTTCTGATCCGAGATG 2580
Db 2562 CTGACCTTCGGGGGCAAGCGGAGTCAATCATCTGTGGCCGCAAGTTCTGATCCGAGATG 2621
QY 2581 GAGAAAGTGGTTGAGAGCAATCCAGATGCGCATTTGACTTGCGGAGAGAGAGAGAGAGAGAG 2640
Db 2622 GAGAAAGTGGTTGAGAGCAATCCAGATGCGCATTTGACTTGCGGAGAGAGAGAGAGAGAGAG 2681
QY 2641 GCCCTGAGTTTCTGCGCAGAGAGCCCTGAGCAACAAGTCCCTGATGAAGCAACCGGG 2700
Db 2682 GCCCTGAGTTTCTGCGCAGAGAGCCCTGAGCAACAAGTCCCTGATGAAGCAACCGGG 2741

QY	2701	GCTGACCAAGAGTCAAGAAATGACCTGAGCGCTTCGGCAATCGCTGGAAGCGCAAGCC	2760
Db	2742	GCTGACCAAGAGTCAAGAAATGACCTGAGCGCTTCGGCAATCGCTGGAAGCGCAAGCC	2801
QY	2761	CCGCAACCGCGGCAACACAATGGTGCACCTGTGCTGGCAACCGCAACACAGCGCTCCATG	2820
Db	2802	CCGCAACCGCGGCAACACAATGGTGCACCTGTGCTGGCAACCGCAACACAGCGCTCCATG	2861
QY	2821	GTGGAATTACGATCGCAGTGGAGAAATAGTTGTGGAAAATTGCTGAGAAATTCAA	2880
Db	2862	GTGGAATTACGATCGCAGTGGAGAAATAGTTGTGGAAAATTGCTGAGAAATTCAA	2921
QY	2881	AACAGCAACCGGTGGCAGAACTGTGGGTGGTTCAGAACTTCTGCTCTTCTTAC	2940
Db	2922	AACAGCAACCGGTGGCAGAACTGTGGGTGGTTCAGAACTTCTGCTCTTCTTAC	2961
QY	2941	AAATCACACAGCACAATCATCCCTTTCACAGCCTGCTCTGCTCGCTACCTCGCTCAC	3000
Db	2982	AAATCACACAGCACAATCATCCCTTTCACAGCCTGCTCTGCTCGCTACCTCGCTCAC	3041
QY	3001	ATCCCTCTGAGTCCGAGAACATCCAGAAAGACTACGTTCAGAGCTCACTCAAGTCC	3060
Db	3042	ATCCCTCTGAGTCCGAGAACATCCAGAAAGACTACGTTCAGAGCTCACTCAAGTCC	3101
QY	3061	CACGCTCTCTACTTTCAGGGCGGAAAGCGAATACAGTTCCGAAAGTGGATGGAAATGATC	3120
Db	3102	CACGCTCTCTACTTTCAGGGCGGAAAGCGAATACAGTTCCGAAAGTGGATGGAAATGATC	3161
QY	3121	CGCAGTGCACACAGCTCTGCTCCGCGAACCCACAGTGTGAGCGACCAAGAGTCTTGTG	3180
Db	3162	CGCAGTGCACACAGCTCTGCTCCGCGAACCCACAGTGTGAGCGACCAAGAGTCTTGTG	3221
QY	3181	TATTGATGGCCGGAACAACATCGTTTCCGAGTGGCTGCTTTCCTGGAAGAGTTTCTTT	3240
Db	3222	TATTGATGGCCGGAACAACATCGTTTCCGAGTGGCTGCTTTCCTGGAAGAGTTTCTTT	3281
QY	3241	CTTCTGTATTATGAAAGCTGTGTAATAATTAACACTGTCTGTAATAATCAAAAACATGCTT	3300
Db	3282	CTTCTGTATTATGAAAGCTGTGTAATAATTAACACTGTCTGTAATAATCAAAAACATGCTT	3341
QY	3301	CCGAGAGAGCTCTCCGCTGTCCACAGCGCGGTTTTTAACTCCCGAAGCTCTGAGCGTTTGA	3360
Db	3342	CCGAGAGAGCTCTCCGCTGTCCACAGCGCGGTTTTTAACTCCCGAAGCTCTGAGCGTTTGA	3401
QY	3361	TGAACAGCGCTCCACACTCGACCTCGTGCACTCCGCTGGGGCGCTGTTCTTTAGCTAGTG	3420
Db	3402	TGAACAGCGCTCCACACTCGACCTCGTGCACTCCGCTGGGGCGCTGTTCTTTAGCTAGTG	3461
QY	3421	CCAGTATTAAACATTTGTCATT	3442
Db	3462	CCAGTATTAAACATTTGTCATT	3483
RESULT 5			
LOCUS	BC041595		
DEFINITION	Homo sapiens FERK, RhogEP (ARHGEF) and pleckstrin domain protein 1 (Chondrocyte-derived), transcript variant 1, mRNA (cDNA clone MGC:55223 IMAGE:5443230), complete cds.		
ACCESSION	BC041595		
VERSION	BC041595.1		
KEYWORDS	MG.		
SOURCE	Homo sapiens (human)		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.		
REFERENCE	1 (bases 1 to 3546)		
AUTHORS	Collinsberg RL, Feinsgold EA, Grouse LH, Derge JG, Klausner RD, Strassburg RS, Wagner L, Shennan CM, Schuler GD, Altschul SF, Zeeberg B, Buetow KH, Schaefer CF, Bhat NK, Hopkins RF, Jordan H, Moore T, Wax ST, Wang J, Hsieh F, Diatchenko L, Marisela K, Farmer AA, Rubinfeld GM, Hong L, Stepleton M, Soares MB, Donald MF, Casavant TL.		

CONSRMT TITLE
 Mammalian Gene Collection Program Team
 Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 12477932
 2 (bases 1 to 3546)
 Director MGC Project.
 Direct Submission
 Submitted (20-DEC-2002) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 REMARK
 CONTACT: MGC help desk.
 NHI-MGC Project URL: <http://mgc.ncl.nih.gov>

Email: cgabds-rc@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland!
Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgc@nhgri.nih.gov
Ahtker, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
Blakeley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
Dietch, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Larric, P., Legaspi, R.,
Maduro, Q.L., Masiello, C., Masketi, B., Mestrian, S.D., McCloskey, J.C.,
McGowan, J., Pearson, R., Scatturipo, S., Thomas, P.J., Touchman, J.W.,
Taurgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNLN at: <http://image.lnl.nih.gov>
Series: IRAL Plate: 44 Row: 1 Column: 24.
Location/Qualifiers
1..3546

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/cclone="MGC:52223 IMAGE:5443210"
/tissue_type="Brain, astrocytoma, grade IV"
/cclone_id="NIH_MGC_98"
/lab_host="DH10B"
/note="Vector: pOTB7"
1..3546
/gene="FARP1"
/gene="FARP1"
/note="synonyms: CDBP, MGC87400"
/db_xref="GeneID:10160"
/db_xref="MIM:602654"
266..3403
CDBP

```

```

/gene="FARP1"
/codon_start=1
/product="FERM, RhogEF, and plectstrin domain protein 1,
isoform 1"
/protein_id="AA41595.1"
/db_xref="GI:49698405"
/db_xref="GeneID:10160"
/db_xref="MIM:602654"
/translation="MGIEORPTPSRGLAPENSGISTLERGQPEPTPSGKLVSIKI
OMLDIDGEAFEPORABGKTLIDAVCNHLNIVGDFGEGFPDHKTIITWLDLKPIV
KQIRKHHVYKVFPPPDHQLOBELIRYIPALQYKODLADGSLTDCSTAAII
SHVQSTIGFDEALDEKTKYIFQDADLDEKIVERRHHITGGTPASSDPQLEI

```


ARLEMYGIRLHPADREBGTKINLAVANTGILVFOGFTKINAFNNAVKRLSPKRRP
LIKRPDANSAVODTLEFLMASRDFCKSPMKI CVENHAFRLFEPEPKPXLVLSRG
SFRSPRQKOVLDYVKEGKHKVOFEKXKIKHSIRSLAFOPLINESEVLEOSQS
TSLTGBRASEGSGOSCRGKRPXVSAGPSGSHSPAPRSPAGNKODGASATPEE
BEVYKDRTOQSKPOPPDSTSLGSPHLSLSPNSOGGVA PAAYTSLSPNSPTKO
ASPLISPLINDQACPTDDEGRRRRFPETDKAYFLAEVSTETRYLKDVLITSWF
OSTVEDKEDMAPLKSILFPNEPLHKFTNPLKEI BQMLAMEGSHNOQIYLDYRIG
DVLKNIQGMKHLAHLIMGSEBALBENGIKSRRLEPCADPELOKVCYPLNTFL
LRLHRLMYKQVLERLCGHPSPHADPCGALAEITBEMVLAOHTGMITKMENTOKL
HEIKDLIGIDLJVGPREFRILGSLSKSGLOORMFLENDULIATYSRLTASNO
FRVHQDPLTYGMITRESEDEKVPKCLTIRGQROSTIYASRSREKKEVEDIQALID
LAEKSSPAEVPFCLASSPDPNKS PDEXTAADQSEEDLSRTSLERQAFHRSNTAVH
CMHRTYSVMDFSI AVENQLSGNLKRKNSNGWOKM VVETFEGLPEFKSHONPH
LASLPLGSLTIPSESENIQKDYVFKLHFKSHVYVYFRAESEYTERMEVIRSATSS
ASRPHVLSHESLIVY"

ORIGIN

Query Match 96.2%; Score 3311; DB 5; Length 3546;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 3311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 CGCCCGAGCGCGCGCGCTGTGGAGATATTCTTCAAGCGCTTTCATCATGGAGAAATA 60
DB |||
QY 218 CGCCCGAGCGCGCGCGCTGTGGAGATATTCTTCAAGCGCTTTCATCATGGAGAAATA 277
DB |||
QY 61 GAGCAGAGCGCGAGCCCGAGATCA CGACTGGGGGCCCGGAAAATTGCGGAGATCAGTACC 120
DB |||
QY 278 GAGCAGAGCGCGAGCCCGAGATCA CGACTGGGGGCCCGGAAAATTGCGGAGATCAGTACC 337
DB |||
QY 121 TTGGAACTGTGACAGAAAGCGCGCCCAACACTTTCAGGAAAATCGTGTTCATCAAAATC 180
DB |||
QY 338 TTGGAACTGTGACAGAAAGCGCGCCCAACACTTTCAGGAAAATCGTGTTCATCAAAATC 397
DB |||
QY 181 CAGATGCTGGATGACACCCAGAGAGCATTTGAAGTTCCAAAGAGCTCTGGGAAAGTGC 240
DB |||
QY 398 CAGATGCTGGATGACACCCAGAGAGCATTTGAAGTTCCAAAGAGCTCTGGGAAAGTGC 457
DB |||
QY 241 CTGCTGGATGCAATTTGCAACAACCTCAACCTCTGTGGAAGTGA CTATTTGGCCTCGAG 300
DB |||
QY 458 CTGCTGGATGCAATTTGCAACAACCTCAACCTCTGTGGAAGTGA CTATTTGGCCTCGAG 517
DB |||
QY 301 TTTCTGATCACAAAAAGATCAACGCTGTGGCTGTGATCTCTTAAACCCATTGTGAACAG 360
DB |||
QY 518 TTTCTGATCACAAAAAGATCAACGCTGTGGCTGTGATCTCTTAAACCCATTGTGAACAG 577
DB |||
QY 361 ATTGAAGGCGCAAGAGCATGTGTTGAAGTTGTGAGTGAATTTCTTCCGCTGACAC 420
DB |||
QY 578 ATTGAAGGCGCAAGAGCATGTGTTGAAGTTGTGAGTGAATTTCTTCCGCTGACAC 637
DB |||
QY 421 ACACAACTCCAAAGAACTCAACAGGTACTGTTCGCGCTGACAGGTGAAGCAGACTTG 480
DB |||
QY 638 ACACAACTCCAAAGAACTCAACAGGTACTGTTCGCGCTGACAGGTGAAGCAGACTTG 697
DB |||
QY 481 GCTCAAGGCAAGTTGACGCTGTATGACACACGCGAGCTCTTGTATTTCAACATTTG 540
DB |||
QY 698 GCTCAAGGCAAGTTGACGCTGTATGACACACGCGAGCTCTTGTATTTCAACATTTG 757
DB |||
QY 541 CAATCGAGATTGGGAGTTTGTATGAAGCTTGAAGAGGAGAGAGAGAGAGAGAGAGAGAG 600
DB |||
QY 758 CAATCGAGATTGGGAGTTTGTATGAAGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 817
DB |||
QY 601 TACATACCTCAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
DB |||
QY 818 TACATACCTCAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 877
DB |||
QY 661 GAGCAAAACACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
DB |||
QY 878 GAGCAAAACACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 937
DB |||
QY 721 TATGGAATCGGAGTGAAGCGCGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
DB |||
QY 938 TATGGAATCGGAGTGAAGCGCGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 997
```

```
QY 781 GCCAACACGGGAATTCAGTGTGTTTCAGGAGTTTCACTAAGATCAATGCTTCAACTGGGCC 840
DB |||
QY 998 GCCAACACGGGAATTCAGTGTGTTTCAGGAGTTTCACTAAGATCAATGCTTCAACTGGGCC 1057
DB |||
QY 841 AAGGTGCGGAAGCTTCAAGTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
DB |||
QY 1058 AAGGTGCGGAAGCTTCAAGTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1117
DB |||
QY 901 AATAGTGGCTACAGAGATACCTTGAATTTCCGATGAGCGAGTCGGGATTTTGCAGAGTCC 960
DB |||
QY 1118 AATAGTGGCTACAGAGATACCTTGAATTTCCGATGAGCGAGTCGGGATTTTGCAGAGTCC 1177
DB |||
QY 961 TTCTGAAAATATGTGTTGAACATCATGCTTTCTTGAACCTTTTGAAGAGCCCAACCA 1020
DB |||
QY 1178 TTCTGAAAATATGTGTTGAACATCATGCTTTCTTGAACCTTTTGAAGAGCCCAACCA 1237
DB |||
QY 1021 AAGCCCAAGCCCGTCTTTTACCGGGGGGTATCATATTCGGTTCAGTGTGAGCTCAG 1080
DB |||
QY 1238 AAGCCCAAGCCCGTCTTTTACCGGGGGGTATCATATTCGGTTCAGTGTGAGCTCAG 1297
DB |||
QY 1081 AAGCAGTTCTCGACTATGTTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
DB |||
QY 1298 AAGCAGTTCTCGACTATGTTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1357
DB |||
QY 1141 CACAGCAAGATTCATTTTACCGAGAGCTTGTCTTCAACAGCTTCAACAGACTGAATTCGAA 1200
DB |||
QY 1358 CACAGCAAGATTCATTTTACCGAGAGCTTGTCTTCAACAGCTTCAACAGACTGAATTCGAA 1417
DB |||
QY 1201 GTGCTGAGAGAGTCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
DB |||
QY 1418 GTGCTGAGAGAGTCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1477
DB |||
QY 1261 GGGGGCCAGAGCTGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
DB |||
QY 1478 GGGGGCCAGAGCTGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1537
DB |||
QY 1321 CACCCGAGCTTGTGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
DB |||
QY 1538 CACCCGAGCTTGTGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1597
DB |||
QY 1381 TGGGGGCCACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
DB |||
QY 1598 TGGGGGCCACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1657
DB |||
QY 1441 CAGCCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500
DB |||
QY 1658 CAGCCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1717
DB |||
QY 1501 GTGAATCTGACAGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1560
DB |||
QY 1718 GTGAATCTGACAGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1777
DB |||
QY 1561 GACACCAAGCAGAGCTTCCCTTGAATCAGCCGCTGTAAGAGAGAGAGAGAGAGAGAGAGAG 1620
DB |||
QY 1778 GACACCAAGCAGAGCTTCCCTTGAATCAGCCGCTGTAAGAGAGAGAGAGAGAGAGAGAGAG 1837
DB |||
QY 1621 ACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1680
DB |||
QY 1838 ACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1897
DB |||
QY 1681 GCTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1740
DB |||
QY 1898 GCTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1957
DB |||
QY 1741 TGGTTTCAGAGCAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1800
DB |||
QY 1958 TGGTTTCAGAGCAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2017
DB |||
QY 1801 TTTCCGAATTTTGAACCTTTGACAAATTTCTAATTAATTTCTCAAGAGAAATTTGAGCAA 1860
DB |||
QY 2018 TTTCCGAATTTTGAACCTTTGACAAATTTCTAATTAATTTCTCAAGAGAAATTTGAGCAA 2077
DB |||
QY 1861 CGACTTGGCTGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1920
```


Db 2078 CGACTTGGCCCTGGGAGGAGCCGCTCAAAATGCCCAATCAGAGATTACAAAAGATGCG 2137
 Qy 1921 GATGTCATGTTGAAGAACTTCAAGGCGATGAAGCACTGCGGCTCACTGTGAAAGCAC 1980
 Db 2138 GATGTCATGTTGAAGAACTTCAAGGCGATGAAGCACTGCGGCTCACTGTGAAAGCAC 2197
 Qy 1981 AGCGAGGCTTGGAGGCGCTGGAGAAATGGAATCAAGAGCTCCGCGGCGCTGGAGAACTTC 2040
 Db 2198 AGCGAGGCTTGGAGGCGCTGGAGAAATGGAATCAAGAGCTCCGCGGCGCTGGAGAACTTC 2257
 Qy 2041 TGCAGAGACTTTGAGCTGAGAGAGTGTGTACTCAACCGCTCAACACTTCTCTGCGG 2100
 Db 2258 TGCAGAGACTTTGAGCTGAGAGAGTGTGTACTCAACCGCTCAACACTTCTCTGCGG 2317
 Qy 2101 CCACTGCAACCGGCTCATGCACTCAACAGAGTCTTGAGAGCGGCTGTGCAAAACACCCG 2160
 Db 2318 CCACTGCAACCGGCTCATGCACTCAACAGAGTCTTGAGAGCGGCTGTGCAAAACACCCG 2377
 Qy 2161 CCGAGGCGCGGCTTCAAGGGAATGCGAGCGGCTTTGGCAGAGATCAAGAGATGATG 2220
 Db 2378 CCGAGGCGCGGCTTCAAGGGAATGCGAGCGGCTTTGGCAGAGATCAAGAGATGATG 2437
 Qy 2221 GCACAGCTCCACGCTACGATGATCAAGATGAGAAATTTCCAGAAAGCTGACGAATCAAG 2280
 Db 2438 GCACAGCTCCACGCTACGATGATCAAGATGAGAAATTTCCAGAAAGCTGACGAATCAAG 2497
 Qy 2281 AAAAGATTGATTGGCATTTGACAACTCTTGTGTTCCGGGAAAGGAGTTCACTCGCTGGGC 2340
 Db 2498 AAAAGATTGATTGGCATTTGACAACTCTTGTGTTCCGGGAAAGGAGTTCACTCGCTGGGC 2557
 Qy 2341 AGCGTCAGAGAGCTCCGGGAAAGGAGCTCCAGAGCGGCTTCTTCTCTGTTCAACGAC 2400
 Db 2558 AGCGTCAGAGAGCTCCGGGAAAGGAGCTCCAGAGCGGCTTCTTCTCTGTTCAACGAC 2617
 Qy 2401 GTCTGCTATTACAGAGCGGCGGCTGACGGCTCCCAATCAGTTTAAAGTCCAGGGGACG 2460
 Db 2618 GTCTGCTATTACAGAGCGGCGGCTGACGGCTCCCAATCAGTTTAAAGTCCAGGGGACG 2677
 Qy 2461 GTCTGCTATTACAGAGCGGCGGCTGACGGCTCCCAATCAGTTTAAAGTCCAGGGGACG 2520
 Db 2678 GTCTGCTATTACAGAGCGGCGGCTGACGGCTCCCAATCAGTTTAAAGTCCAGGGGACG 2737
 Qy 2521 CTGACCCCTCCGGGCGCAGGCGGCTCCATCATGTCGCGGCTTCTGCTCCGAGATG 2580
 Db 2738 CTGACCCCTCCGGGCGCAGGCGGCTCCATCATGTCGCGGCTTCTGCTCCGAGATG 2797
 Qy 2581 GAGAAAGTGGTTGAGAGACATCCAGATGCGCATTTGACCTGGCGGAGAAAGACAGACGCC 2640
 Db 2798 GAGAAAGTGGTTGAGAGACATCCAGATGCGCATTTGACCTGGCGGAGAAAGACAGACGCC 2857
 Qy 2641 GCCCCTGAGTTCTCTGGCCAGCAGCGCCCTCTGACCAAGAGTCCCTGATGAAAGCCAGCGG 2700
 Db 2858 GCCCCTGAGTTCTCTGGCCAGCAGCGCCCTCTGACCAAGAGTCCCTGATGAAAGCCAGCGG 2917
 Qy 2701 GGTGACACAGAGTCCAGAGATGACCTGACCGCTCCCGCACTCGCTGAGAGCGCCAGGCC 2760
 Db 2918 GGTGACACAGAGTCCAGAGATGACCTGAGCGCTCCCGCACTCGCTGAGAGCGCCAGGCC 2977
 Qy 2761 CCGCAGCGCGGCAACCAATGTCAGCTGTGCGACCGCAACACAGAGCTCTCATG 2820
 Db 2978 CCGCAGCGCGGCAACCAATGTCAGCTGTGCGACCGCAACACAGAGCTCTCATG 3037
 Qy 2821 GTGAGACTTCAAGTCCAGATGAGAAATCAAGTTCTGTGAAACTGCTGAGAAATTCAAA 2880
 Db 3038 GTGAGACTTCAAGTCCAGATGAGAAATCAAGTTCTGTGAAACTGCTGAGAAATTCAAA 3097
 Qy 2881 AACAGCAAGGGGAGGAGAGCTGTGGGTGTCTCAAACTTCTGCGCTCTCTCTAC 2940
 Db 3098 AACAGCAAGGGGAGGAGAGCTGTGGGTGTCTCAAACTTCTGCGCTCTCTCTAC 3157
 Qy 2941 AATATCACACAGAGCAATCATCCCTTGGCAGCTCTCTGCTCGGCTACTGCTACAC 3000

Db 3158 AATATCACACAGAGCAATCATCCCTTGGCAGCGCTGCTGCTCGGCTACTGCTCAC 3217
 Qy 3001 ATCCCTCTGAGTCCGAGAACTCCAGAAAGACTAGTGTTCAGAGCTTCAAGTCC 3060
 Db 3218 ATCCCTCTGAGTCCGAGAACTCCAGAAAGACTAGTGTTCAGAGCTTCAAGTCC 3277
 Qy 3061 CAGCTCACTACTTCAAGGCGGAGAGGAGTACAGCTTGGAAAGGTGATGATGATC 3120
 Db 3278 CAGCTCACTACTTCAAGGCGGAGAGGAGTACAGCTTGGAAAGGTGATGATGATC 3337
 Qy 3121 CCGAGTCCACCAAGCTCTGCTCGCGACCCCAAGTGTGAGCCACAAAGAGTCTTTGTG 3180
 Db 3338 CCGAGTCCACCAAGCTCTGCTCGCGACCCCAAGTGTGAGCCACAAAGAGTCTTTGTG 3397
 Qy 3181 TATTGATGCGCGGACACACTGCTTTCCGAGTGGCTTTCTTCTGGAAGCTTTCTTT 3240
 Db 3398 TATTGATGCGCGGACACACTGCTTTCCGAGTGGCTTTCTTCTGGAAGCTTTCTTT 3457
 Qy 3241 CTTCTGATTAATGAAGCCGTGTAATAATTAACACTGTCTGAATAATCAAAATCATGCTT 3300
 Db 3458 CTTCTGATTAATGAAGCCGTGTAATAATTAACACTGTCTGAATAATCAAAATCATGCTT 3517
 Qy 3301 CCCAGCAGCTC 3311
 Db 3518 CCCAGCAGCTC 3528

RESULT 6
 BC071592
 LOCUS
 DEFINITION
 Homo sapiens PERM, RHOGEF (ARHGEF) and pleckstrin domain protein 1 (chondrocyte-derived), mRNA (cDNA clone MGC:87400 IMAGE:30342239), complete cds.
 BC071592
 BC071592.1 GI:47940451
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 4932)
 REFERENCES
 AUTHORS
 Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, P.S., Wagner, L., Shennan, C.M., Schlier, G.D., Altschuld, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Uedl, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loggellano, N.A., Peters, G.J., Abramson, R.D., Muliyil, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Huiyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahy, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shcherchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butlerfield, Y.S., Krzywinski, M.I., Skalska, U., Smallus, D.E., Scherch, A., Schein, J.B., Jones, S.J. and Marra, M.A.
 Mammalian Gene Collection Program Team
 Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 12477932
 2 (bases 1 to 4932)
 JOURNAL
 PUBLISHED
 REFERENCE
 AUTHORS
 CONSRM
 TITLE
 CONSRM
 JOURNAL
 REMARK
 COMMENT
 NIH MGC Project
 Direct Submission
 Submitted (01-JUN-2004) National Institutes of Health, Mammalian Gene Collection (MGC), Bethesda, MD 20892-2590, USA
 NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
 Contact: MGC help desk

Email: cgaabs-r@mail.nih.gov
 Tissue Procurement: Dr. Stefan Hansson
 cDNA Library Preparation: Michael Brownstein / Ted Udén
 Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome
 Center, Stanford University School of Medicine, Stanford, CA 94305
 Web site: <http://www-shgc.stanford.edu>
 Contact: (Dickson, Mark) mdickpaxil.stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
 R. M.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>
 Series: IRAX Plate: 168 Row: 1 Column: 8
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 48528036.

FEATURES

source

1..4932
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:87400 IMAGE:30342239"
 /issue_type="Placenta, normal"
 /clone_1b="NIH_MGC_147"
 /lab_host="DH10B"
 /note="Vector: pBluescriptR"

gene

1..4932
 /gene="FARP1"
 /note="synonyms: CDEP, MGC87400, PLEKHC2"
 /db_xref="GeneID:10160"
 /db_xref="MIM:602654"

CDS

240..3470
 /gene="FARP1"
 /codon_start=1
 /product="FARP1 protein"
 /protein_id="AAH71592.1"
 /db_xref="GI:47940452"
 /db_xref="GeneID:10160"
 /db_xref="MIM:602654"
 /translation="MGEIQRPTPGSRGLAPENSGISTLEBROKPPPTPSGLVSLIKI
 QMLDPTQEAFFVQPAHGVKVLIDAVCNHLNLTVEGDYFGLSPDHKKITVWLKPLIV
 KQIRPKVHVVKVFPFDDHLOLELTRVFLQVQDLAQGLTNDTSALLI
 SHIVSEIDFDEALDRHLAKNKYIQDDALEDKIVEHNHNIIGTAPESFOLEI
 ARLEWYGRILHAPAKDRBGKINLANVANGILVFOGFTINFNNAKVKLSFKRRF
 LTKAPDANSAYODTLEPLMASRDECFKFWKI CENHAFPLFEBPKPKPVLRSRG
 SFPSPSGRTQKQVLDVKEGKHQVQFEKHSKHSIRLSAQPTLEINSEVLEQOQS
 TSLTVEGASFPQGSCKRGKPKVSAGPQSHSPAPKRSFPAKNSQADGASAPTEE
 EBEVYDRTOQSPQPOPSTGLTGSPLSLVNSQGVAPAVNTLSPNLSPTIKO
 ASPLISPLINDQACPTDEDEGRKRKPTDRAVFAKESSTTERVYLKDLIVITWF
 OSTVSEKDEMPALKSLIPNPEPLHKPTNPLKEIQRALMEGRSNQOISDYORIG
 DMLKNIQGMKHLAHLMKYSRLEALENGISRSRLRNFRCDFPELOKVCYLLPMTFL
 LRLRLHMYTKQVLERLCKHHPSHADPFDCRAALAEITBEMVQLHGMTIKENOKL
 HELKDLIGIDNVVGRPGFSVLMKTPHLAGHRRPCAPERPLLVKFTRLSLS
 KLSGKLOQRMFFLPNDVLLYTSRGLTASNQPKVHQQLPYGMTIEESPDENVPHCL
 TLRSGRQSIIVAAFSRSEMEKWEVDIQMIDLAEKSSLPAPFLIASPDKNSPEBAT
 AADQSEBDLSASRTSLERQAPHRGNTVCHMNTVSMTDFSIAYENQSGMLIK
 KKNNGMOKLMVFTNPLCYKSHQNDHPLASLPGLGSLTISESNIQKDYVFK
 LHFKSHVYFRASEETTFERMMEVIRSAISASRPVLSHKESLYV"

ORIGIN

Query Match 96.2%; Score 3310.2; DB 5; Length 4932;
 Best Local Similarity 97.3%; Pred. No. 0;
 Matches 3415; Conservative 0; Mismatches 3; Indels 93; Gaps 1;

QY 25 GATATTCTTAAAGCCGCTTCATCATGGAGAAATGAGACAGGCCGACCCAGATCA 84
 |||
 DB 216 GATATTCTTAAAGCCGCTTCATCATGGAGAAATGAGACAGGCCGACCCAGATCA 275
 |||
 QY 85 CGACTGGGGGCCCCGGAATAATCGGGGATCAGTACCTTGAACGTTGAGACAGAACGCCGCC 144
 |||
 DB 276 CGACTGGGGGCCCCGGAATAATCGGGGATCAGTACCTTGAACGTTGAGACAGAACGCCGCC 335
 |||

QY 145 CCAACACCTTCAGAGAAAATCGTGTCATCAAAATTCAGANTCTGATGACACCCAGAG 204
 |||
 DB 336 CCAACACCTTCAGAGAAAATCGTGTCATCAAAATTCAGANTCTGATGACACCCAGAG 395
 |||
 QY 205 GATTTTGAAGTTCCCAAAAGAGCTCTGAGAAAGTGTCTGTGATGATTTTGAACCCAC 264
 |||
 DB 396 GATTTGAAGTTCCCAAAAGAGCTCTGAGAAAGTGTCTGTGATGATTTTGAACCCAC 455
 |||
 QY 265 CTCACACTCGTGGAGAGTGACTATTGTTGGCCCTGAGATTCTGATCACAAAAGATCAAG 324
 |||
 DB 456 CTCACACTCGTGGAGAGTGACTATTGTTGGCCCTGAGATTCTGATCACAAAAGATCAAG 515
 |||
 QY 335 GTTGGCTGGATCTCTTAAACCCATTGTGAACAGATTGAAGGCCAAAGACCTGTGTT 384
 |||
 DB 516 GTTGGCTGGATCTCTTAAACCCATTGTGAACAGATTGAAGGCCAAAGACCTGTGTT 575
 |||
 QY 385 GTTAAGTTTGTGTGAATTTCTTCCGCTGACACACAACTCCAGAAAGAACCTCAC 444
 |||
 DB 576 GTTAAGTTTGTGTGAATTTCTTCCGCTGACACACAACTCCAGAAAGAACCTCAC 635
 |||
 QY 445 AGGTACTGTTGCGCTGAGGTGAAGCAGACTTGGCTCAAGCAGGTTGACGTGTAT 504
 |||
 DB 636 AGGTACTGTTGCGCTGAGGTGAAGCAGACTTGGCTCAAGCAGGTTGACGTGTAT 695
 |||
 QY 505 GACACACGAGCAGCTCTCTTGAATTTACACATTTGTGCAATCTGAGATTGGGGAATTTGAT 564
 |||
 DB 696 GACACACGAGCAGCTCTCTTGAATTTACACATTTGTGCAATCTGAGATTGGGGAATTTGAT 755
 |||
 QY 565 GAAGCTTGGAGACAGAGACACTTACCAAAAATAATACATACCTCAGCAAGACGACAT 624
 |||
 DB 756 GAAGCTTGGAGACAGAGACACTTACCAAAAATAATACATACCTCAGCAAGACGACAT 815
 |||
 QY 625 GAGACAAAATCGTGAATTTTACCATTAACCATTTGACAAACACACAGACAAATTCAGT 684
 |||
 DB 816 GAGACAAAATCGTGAATTTTACCATTAACCATTTGACAAACACACAGACAAATTCAGT 875
 |||
 QY 685 TTCCAGCTCTTGAAGATGTCCTGCTGCTGAGATGATGAAATCCGTTGCAACCCGCC 744
 |||
 DB 876 TTCCAGCTCTTGAAGATGTCCTGCTGCTGAGATGATGAAATCCGTTGCAACCCGCC 935
 |||
 QY 745 AAGACAGAGGAAGGACCAAGATCAATCGCGCTTCCCAACGGAATTCATGAGTTT 804
 |||
 DB 936 AAGACAGAGGAAGGACCAAGATCAATCGCGCTTCCCAACGGAATTCATGAGTTT 995
 |||
 QY 805 CAGGTTTCACTAATGATCAATCTCTTCAATCGGACCAAGTGCAGAGCTTCAAG 864
 |||
 DB 996 CAGGTTTCACTAATGATCAATCTCTTCAATCGGACCAAGTGCAGAGCTTCAAG 1055
 |||
 QY 865 AGGAAGCGCTTCTCATCAAGCTCCGGCCAGATGCCAATAGTGCCTACAGATACCTTG 924
 |||
 DB 1056 AGGAAGCGCTTCTCATCAAGCTCCGGCCAGATGCCAATAGTGCCTACAGATACCTTG 1115
 |||
 QY 925 GAATTCCTGATGCGCAGTCGGGATTTTCAGAGTCTCTGAAAAATCTGTGTGAACAT 984
 |||
 DB 1116 GAATTCCTGATGCGCAGTCGGGATTTTCAGAGTCTCTGAAAAATCTGTGTGAACAT 1175
 |||
 QY 985 CATGCTCTTTAGACTTTTGAAGAGCCCAACCAAGCCCAAGCCCTCTCTTAAAC 1044
 |||
 DB 1176 CATGCTCTTTAGACTTTTGAAGAGCCCAACCAAGCCCAAGCCCTCTCTTAAAC 1235
 |||
 QY 1045 CGGGGGCTCATATTTGCTTCAAGTGTGAGACTCAAGACAGGTTCTGACTATGTTAA 1104
 |||
 DB 1236 CGGGGGCTCATATTTGCTTCAAGTGTGAGACTCAAGACAGGTTCTGACTATGTTAA 1295
 |||
 QY 1105 GAAGAGGACATTAAGAGTGTGAGTTTGAAGAGGACAGAAAGTTGATTCATCCGG 1164
 |||
 DB 1296 GAAGAGGACATTAAGAGTGTGAGTTTGAAGAGGACAGAAAGTTGATTCATCCGG 1355
 |||
 QY 1165 AGCTTGTCTTCAACACCTTACAGAACTGAATTCGGAAGTCTGAGACGTTCTCAGCAGAG 1224
 |||
 DB 1356 AGCTTGTCTTCAACACCTTACAGAACTGAATTCGGAAGTCTGAGACGTTCTCAGCAGAG 1415
 |||
 QY 1225 ACCAGCTTACATTTTGAAGAGTGTCCGAATCTCCAGGGGGCCAGAGCTCCGGCGAGGA 1284
 |||

```
Db 1416 ACAGAGCTTACATTTGGAGAGAGGTGCGGATCTCCAGGGGGCCAGAGCTCCGGCCAGAGA 1475
Qy 1285 AAGGAACCGAAGGTTTCCCGCGGGAGCCGGGGTCCGACCCGAGCCCTCGCCGAGAGA 1344
Db 1476 AAGGAACCGAAGGTTTCCCGCGGGAGCCGGGGTCCGACCCGAGCCCTCGCCGAGAGA 1535
Qy 1345 AGCCCCGGGGTTAAACAGAGAGGGGAGCCGAGCCGGCCCTCGGGCCGACCGAGGAAGAGAG 1404
Db 1536 AGCCCCGGGGTTAAACAGAGAGGGGAGCCGAGCCGGCCCTCGGGCCGACCGAGGAAGAGAG 1595
Qy 1405 GAGGTGTTAAGGATAGAGCCAGAGAGTAACTTCAGCCCGCCGAGCCGACCAAGAGC 1464
Db 1596 GAGGTGTTAAGGATAGAGCCAGAGAGTAACTTCAGCCCGCCGAGCCGACCAAGAGC 1655
Qy 1465 TCCCTGACTGGAGCTCTCACTTTCCGAGCTGTCTGTAACTCCGAGGGGGAGTGGCC 1524
Db 1656 TCCCTGACTGGAGCTCTCACTTTCCGAGCTGTCTGTAACTCCGAGGGGGAGTGGCC 1715
Qy 1525 CCGGCAAGGTGACTTGTCTCCCACTGAGCCCGAGCCGAGCCGAGAGGCTCTCCCTTG 1584
Db 1716 CCGGCAAGGTGACTTGTCTCCCACTGAGCCCGAGCCGAGCCGAGAGGCTCTCCCTTG 1775
Qy 1585 ATCAGCCCGCTGCTGATGACAGAGCTGCCCCGAGCCGAGCCGATGAGTGAAGGCGG 1644
Db 1776 ATCAGCCCGCTGCTGATGACAGAGCTGCCCCGAGCCGAGCCGATGAGTGAAGGCGG 1835
Qy 1645 AGAAGAGATTTCCCACTGATTAAGCGTACTTCACTAGCTAAGAGAGTGTCTTACCCAG 1704
Db 1836 AGAAGAGATTTCCCACTGATTAAGCGTACTTCACTAGCTAAGAGAGTGTCTTACCCAG 1895
Qy 1705 CGAACAATCTGAGAGATCTCGAAGTTATCACTTCGTGGTTTGAAGCACAGAGAGAA 1764
Db 1896 CGAACAATCTGAGAGATCTCGAAGTTATCACTTCGTGGTTTGAAGCACAGAGAGAA 1955
Qy 1765 GAGAGCGCATGCGGAGAGCACTGAAAGTCTCATATTCGGAATTTTGAACCTTTGAC 1824
Db 1956 GAGAGCGCATGCGGAGAGCACTGAAAGTCTCATATTCGGAATTTTGAACCTTTGAC 2015
Qy 1825 AAATTTCACTAATTTTCTCAAGAAATTGAGCAACGACTTCCCTGTGGAGAGCGCG 1884
Db 2016 AAATTTCACTAATTTTCTCAAGAAATTGAGCAACGACTTCCCTGTGGAGAGCGCG 2075
Qy 1885 TCAAATGCCAATTCAGAGATTACAAAGAAATGGGAGATGTCATGCTGAAGAACTTCAG 1944
Db 2076 TCAAATGCCAATTCAGAGATTACAAAGAAATGGGAGATGTCATGCTGAAGAACTTCAG 2135
Qy 1945 GGCATGAAGCACTGCGGCGCTCACTGTGGAACAGAGAGGCTTTGAGGCGCTGAG 2004
Db 2136 GGCATGAAGCACTGCGGCGCTCACTGTGGAAGTACAGAGAGGCTTTGAGGCGCTGAG 2195
Qy 2005 AATGGAATCAGAGACTCCCGGCGGCTGGAGAACTTCTGAGAGACTTTGAGCTGAGAG 2064
Db 2196 AATGGAATCAGAGACTCCCGGCGGCTGGAGAACTTCTGAGAGACTTTGAGCTGAGAG 2255
Qy 2065 GTGTGTACTTACCGGCTCAACACTTCTCCCTGCGGCACTGACAGGCTCATGCACTAC 2124
Db 2256 GTGTGTACTTACCGGCTCAACACTTCTCTCTGCGGCACTGACAGGCTCATGCACTAC 2315
Qy 2125 AAGCAGGTCTGAGAGGCTGTGCAAAACACACCCCGAGCCGAGCCGACTTCAAGAGAC 2184
Db 2316 AAGCAGGTCTGAGAGGCTGTGCAAAACACACCCCGAGCCGAGCCGACTTCAAGAGAC 2375
Qy 2185 TGCAGAGCGGCTTTGGCAGAGATCAGCGAGATGTGGCACAAGTCTTCCAGGTACGATATC 2244
Db 2376 TGCAGAGCGGCTTTGGCAGAGATCAGCGAGATGTGGCACAAGTCTTCCAGGTACGATATC 2435
Qy 2245 AAGATGAGAAATTTCCAGAACTGCAAGAACTCAAGAAAGATTTGATGGCAATTGCAAT 2304
Db 2436 AAGATGAGAAATTTCCAGAACTGCAAGAACTCAAGAAAGATTTGATGGCAATTGCAAT 2495
Qy 2305 CTTGTGGTTCCGGAAG----- 2321
Db 2496 CTTGTGGTTCCGGAAGGCTGTGCTCTTTTCCCTTATGAGGACTCCACACTTGAAGCAA 2555
Qy 2322 -----GAGTTCATC 2331
Db 2556 GCCAGGCGCATCCCATGTGTCTCCAGAGAGAGGCCATTGTTAATGTAAGAAAGATTCATC 2615
Qy 2332 CGTCTGGGAGGCTTCCAGCAAGCTCTCGGGGAAAGGGGCTCCAGAGCGCATGTTCTCTG 2391
Db 2616 CGTCTGGGAGGCTTCCAGCAAGCTCTCGGGGAAAGGGGCTCCAGAGCGCATGTTCTCTG 2675
Qy 2392 TTCAGCAGCTCTGCTATACAGAGCCGGGGGCTGACGGCTTCAATCACTGTTAAAGTC 2451
Db 2676 TTCAGCAGCTCTGCTATACAGAGCCGGGGGCTGACGGCTTCAATCACTGTTAAAGTC 2735
Qy 2452 CAGGGCAGCTCCCGCTTATATGCAATGACATGACATTTGAAGAGAGCGAAGAGTGGGGTG 2511
Db 2736 CAGGGCAGCTCCCGCTTATATGCAATGACATGACATTTGAAGAGAGCGAAGAGTGGGGTG 2795
Qy 2512 CCCCACTGCTGACCCCTCGGGGCGCAGCGGAGTCCATCATCTGTGGCCGCACTTCTCGG 2571
Db 2796 CCCCACTGCTGACCCCTCGGGGCGCAGCGGAGTCCATCATCTGTGGCCGCACTTCTCGG 2855
Qy 2572 TCCGAGATGAGAAATGAGGTTTGAAGCATTCAGATGCGCATTTGACCTTGGGAGAAAGC 2631
Db 2856 TCCGAGATGAGAAATGAGGTTTGAAGCATTCAGATGCGCATTTGACCTTGGGAGAAAGC 2915
Qy 2632 AGAAGCCCCGCTCTGAGTTCTGTGGCAGAGGCCCCCTGACAAACAAAGTCCCTGATGA 2691
Db 2916 AGAAGCCCCGCTCTGAGTTCTGTGGCAGAGGCCCCCTGACAAACAAAGTCCCTGATGA 2975
Qy 2692 GCCACGCGGCTGACCAAGAGTGAAGATGACCTGAGCGCTTCGCGCATCGCTGAG 2751
Db 2976 GCCACGCGGCTGACCAAGAGTGAAGATGACCTGAGCGCTTCGCGCATCGCTGAG 3035
Qy 2752 CGCAGGCCCCGAGCCGCGGCAACACAAATGTCACAGTGTCTGGCACCGCAACCCAGC 2811
Db 3036 CGCAGGCCCCGAGCCGCGGCAACACAAATGTCACAGTGTCTGGCACCGCAACCCAGC 3095
Qy 2812 GTTCCATGATGGAATTTCAAGATTCGAGATGGAAGATTCAGTGTCTGGAACTGTGAGG 2871
Db 3096 GTTCCATGATGGAATTTCAAGATTCGAGATGGAAGATTCAGTGTCTGGAACTGTGAGG 3155
Qy 2872 AAATTCAAAAACGAGAGGAGTGGAGAGCTGTGGGTGTGTCAAAACTTCTGCTG 2931
Db 3156 AAATTCAAAAACGAGAGGAGTGGAGAGCTGTGGGTGTGTCAAAACTTCTGCTG 3215
Qy 2932 TTCTTTCAAAATCACACAGAGAACTATCCCTTGCAGAGCTGTGCTCGGCTAC 2991
Db 3216 TTCTTTCAAAATCACACAGAGAACTATCCCTTGCAGAGCTGTGCTCGGCTAC 3275
Qy 2992 TCGCTACCATCCCTCTGTAGTCCGAGACATTCAGAAAGATCACTGTTCAGAGCTGAC 3051
Db 3276 TCGCTACCATCCCTCTGTAGTCCGAGACATTCAGAAAGATCACTGTTCAGAGCTGAC 3335
Qy 3052 TTCAAGTCCAGCTTACTTACTTCAAGGCGGAAAGGAGTACAGCTTTCGAAAGTGGATG 3111
Db 3336 TTCAAGTCCAGCTTACTTACTTCAAGGCGGAAAGGAGTACAGCTTTCGAAAGTGGATG 3395
Qy 3112 GAAATGATCCGAGTCCACAGAGTGTGCTCGGACCCCAAGTGTGAGCCACAAAGAG 3171
Db 3456 GAAATGATCCGAGTCCACAGAGTGTGCTCGGACCCCAAGTGTGAGTCAAAAGAG 3455
Qy 3172 TCTCTGTGTATTTGATGAGCGGAGACACATCTGTTCCGCAATGTGCTTCTCTGAAAGC 3231
Db 3456 TCTCTGTGTATTTGATGAGCGGAGACACATCTGTTCCGCAATGTGCTTCTCTGAAAGC 3515
Qy 3232 GTTTCCTTCTTCTGTATTAATGAAGCTGTGTAATTAACACTCTGTGAATAATCAAAA 3291
Db 3516 GTTTCCTTCTTCTGTATTAATGAAGCTGTGTAATTAATTAACACTCTGTGAATAATCAAAA 3575
Qy 3292 ACATGGCTTCCAGAGCTCTCTGTCTCCACAGCGCGTTTTAAACCCCGACCTCTCA 3351
Db 3576 ACATGGCTTCCAGAGCTCTCTGTCTCCACAGCGCGTTTTAAACCCCGACCTCTCA 3635
```

QY 3352 GCGTTTGAATGAACAGCGCTCCACCTCCAGTCTCGCATCCGCTGGGGGCGCTGTTCTT 3411
|||||
Db 3636 GCGTCTGAATGAACAGCGCTCCACCTCCAGTCTCGCATCCGCTGGGGGCGCTGTTCTT 3695
|||||
QY 3412 TAGCTAGTCCAGTATTAAACATTGTCAAT 3442
|||||
Db 3696 TAGCTAGTCCAGTATTAAACATTGTCAAT 3726
|||||
RESULT 7
BC030329 3902 bp mRNA linear ROD 21-OCT-2003
LOCUS Mus musculus FERM, Rhogef (Arhgef) and pleckstrin domain protein 1
DEFINITION (chondrocyte-derived), mRNA (cDNA clone IMAGE:5376197), partial
cds.
BC030329
BC030329.1 GI:20987935
Mus musculus (house mouse)
SOURCE
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridea; Muridae; Murinae; Mus.
1 (bases 1 to 3902)
Strausberg, R.D., Collins, F.S., Wagner, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L.H., Derge, J.G.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diachenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Schertz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
Carinci, P., Prange, C., Raha, S.S., Loggellano, N.A., Peters, G.J.,
Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.D., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kettaman, M., Madan, A., Young, A.C., Shevchenko, Y.,
Sanchez, A., Whiting, M., Madan, A., Touchman, J.W., Green, E.D.,
Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Small, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 3902)
Strausberg, R.
Direct Submission
Submitted (07-MAY-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
Contact: MGC help desk
Email: cgabbs-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgc@nsl.nih.gov
Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
Blakeley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
Dierich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Latic, P., Legaspi, R.,
Maduro, O.L., Mastello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,
McDowell, J., Pearson, R., Stantip, S., Thomas, P.J., Touchman, J.W.,
Tsugeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
Young, A., Zhang, L.-H. and Green, E.D.

FEATURES
source
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
Series: IRAP Plate: 65 Row: m Column: 19
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis.
Location/Qualifiers
1..3902
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N-3"
/db_xref="taxon:10090"
/clone="IMAGE:5376197"
/tissue_type="Mammary tumor, MMTV-LTR/INT3 model, 5 month
old mouse, Taken by biopsy."
/clone_lib="NCI_CGAP_Mam2"
/lab_host="DH10B"
/note="vector: pCMV-SPORT6"
<1..3902
/gene="Parp1"
/note="synonym: Cdep"
/db_xref="geneid:223254"
/db_xref="MGI:246173"
<1..2510
/gene="Parp1"
/product="Parp1 protein"
/protein_id="AAH30329.1"
/db_xref="GI:20987936"
/db_xref="geneid:223254"
/db_xref="MGI:246173"
/translation="TRPEVARRLLEMYGILRLHPAKDREGKINLAVANTGILVFOGFTK
INAPWAVRKLSEPRKFLIKLPDNNSSVODTLEPMAGDPCKSPFKITVEHNAF
RLPEPRKPEPVLPSSGSSRRSGRQKQVLDVYKGGKRVQPERKSKTHRS
LVSGTPANSEVPKSPQASLTLEEGTESPEGQSQQAKETKACTLELGPQSPALP
KSPSPSKADGTVVPPPEEBEEBEGKDGIRSNPQPPQPSGLTSGPHLESLNS
OGGAPAVNTLSPNLSPNKOKASPLISPLNDQACPRDDEBGRKRPPTKAVYIA
KEVSTERTVYLDELIVASIPQSTVSKDSMEALKSLIFPVEPLKHTPTLKIE
ORLALMBRSNAHVGDYORIGDVLKNIOGKHLAAHLMKKSLEALETSIKSRR
LEHFRDLRQVCTLPPLTLRLPLRLHMLRYHLERLCRKHPRNHADPRCCRALA
ETTEVVALDGLTMIWENFOKHEHLKQDILGIDNVTPREBTRIGLSLKSQGGQQ
RNMFLENDVLVLTSGSLTASNOFKHGLPLVGMTEBEEBEGVPHCTLGRQSTI
IYAASSREMEKEMEDIOALDIAKSGPTEPELLASSPPNKSDEATPAODESDI
LSASRTSLERQAPHRGNTVMVCHMRSTSVSNVDSIAVENOLSGNLKAFKNSGMO
KLWVFTNRCLEFFYKSHODSHPLALPLILGSLTTPSEENHKQYVFLRHKSHYY
FRASEEYTERMEVIRSATSSASRAHLISKESHLY"
999..1556
/gene="Parp1"
/note="Rhogef; Region: Guanine nucleotide exchange factor
for Rho/Rac/Cdc42-like GTPases"
/db_xref="CCD:smart00325"
1695..1946
/gene="Parp1"
/note="PH; Region: Pleckstrin homology domain. Domain
commonly found in eukaryotic signalling proteins. The
domain family possesses multiple functions including the
abilities to bind inositol phosphates, and various
proteins. PH domains have been found to possess inserted
domains (such as in PLC gamma, syntrophins) and to be
inserted within other domains. Mutations in Brutons
tyrosine kinase (Btk) within its PH domain cause X-linked
agammaglobulinemia (XLA) in patients. Point mutations
cluster into the positively charged end of the molecule
around the predicted binding site for phosphatidylinositol
lipids"
/db_xref="CCD:smart00233"
2175..2465
/gene="Parp1"
/note="PH; Region: Pleckstrin homology domain. Domain
commonly found in eukaryotic signalling proteins. The
domain family possesses multiple functions including the
abilities to bind inositol phosphates, and various

misc_feature

misc_feature

misc_feature

CDS

gene

proteins. PH domains have been found to possess inserted domains (such as in pLc gamma, syntenophins) and to be inserted within other domains. Mutations in Bruton's tyrosine kinase (Btk) within its PH domain cause X-linked agammaglobulinemia (XLA) in patients. Point mutations cluster into the positively charged end of the molecule around the predicted binding site for phosphatidylinositol 1phide"
/db_xref="CD:smat00233"

ORIGIN

Query Match 56.0%; Score 1927.8; DB 6; Length 3902;
Best Local Similarity 85.9%; Pred. No. 0;
Matches 2165; Conservative 0; Mismatches 347; Indels 9; Gaps 2;

```

QY 690 GCTCCCTAGAGATTCCCGCTGAGATGTATGGAATCCGGTTGACCCGGCCAAAGA 749
DB 5 GCGTCGGAGGTGGCCCGCTGAGATGTATGGAATCCGGTTGACCCCGCCAAAGA 64
QY 750 CAGGGAAGGACGAGATCAATCTGGCCGTGGCAACAGGGAATCTAGTTTCAAGG 809
DB 65 CAGGGAAGGACGAGATCAATCTGGCTGTGGCAACAGGGAATCTAGTTTCAAGG 124
QY 810 TTTCACTAAGATCAATGCTTTCATCTGGCCAAAGTGGAGCTGAGCTTCAAGAGA 869
DB 125 CTTCACGAGATCAATGCTTTCATCTGGCCAAAGTGGAGCTGAGCTTCAAGAGGA 184
QY 870 GGGCTTTCTCATAGACTCCGGCCAGATGCCAATAGTCGTACAGAGATCTTGAAT 929
DB 185 GGGCTTTCTCATAGACTCCGGCCAGATGCCAATAGTCGTACAGAGATCTTGAAT 244
QY 930 CCGATGGCCAGTCCGGATTTCTGCAAGTCTCTGGAATACTGTGTGAATCATATGC 989
DB 245 CCGATGGCCAGTCCGGATTTCTGCAAGTCTCTGGAATACTGTGTGAATCATATGC 304
QY 990 CTTCTTTAGACTTTTGAAGAGCCCAAGCCCAAGCCCGCTCTTTAGCCGGG 1049
DB 305 CTTCTTTAGACTTTTGAAGAGCCCAAGCCCAAGCCCGCTCTTTAGCCGGG 364
QY 1050 GTTCATGATTTCCGTTCACTGCTGAGTCAAGAGAGTTCTGACTATGTAAGAG 1109
DB 365 GTTCATGATTTCCGTTCACTGCTGAGTCAAGAGAGTTCTGACTATGTAAGAG 424
QY 1110 AGGACATTAAGAGTTCAGTTGAAGAGACACAGATTCATTTATCCGAGCCT 1169
DB 425 CCGCCACAAAGAGTTCAGTTGAAGAGACACAGATTCATTTATCCGAGCCT 484
QY 1170 TGCCTCACAGCTTACAGAACTGAATTCGAAAGTGTGAGCACTTCAGAGACCA 1229
DB 485 GGTCTCACAGCTTACAGAACTGAATTCGAAAGTGTGAGCACTTCAGAGACCA 544
QY 1230 CCTTACATTTGAGAAAGTGCAGATCTCCAGGGGCCAGAGCTCCCGGCGAGAAAG 1289
DB 545 CCTTACATTTGAGAAAGTGCAGATCTCCAGGGGCCAGAGCTCCCGGCGAGAAAG 604
QY 1290 ACCGAAAGTTTCCGCGGGGAGCCGGGTCGACCCGAGGCCCTGCCAGAGAAAGCC 1349
DB 605 AACCAAAAGCTTGAAGTTGGAAGTTCACCAAGCCCGCCCTCCCTTAAGAGCC 664
QY 1350 CCGCGGTAAACAAGCAGAGCCGAGCCGCTCCGCGCCCAAGAGAAAGAGAGAGT 1409
DB 665 TCCAGGAGCAAAAGTGCAGATGGAACCAAGTGTGCGGCTCGAGAGAAAGAGAG 724
QY 1410 CGTT-----AAGGATAGAACCAAGAGTAACCTCAGCCCCCGCAGCAAGCAAG 1463
DB 725 AAGAGGTGCAAGAGAGGATCCGGCCAAAGTAACCTCAGCCCCCGCAGCAAGCAAG 784
QY 1464 CTCCCGAGTCCGAGCTTCACTTTCAGAGCTGTGGAAGTCTCGAGGGGGAGAGTGG 1523
DB 785 CTCCCGAGTCCGAGCTTTCAGAGCTGTGGAAGTCTTCACTCAGAGAGAGAGCTGC 844
QY 1524 CCCTGCAACGTGACCTTGTCTTCCAACTGAGCCCGCAGACCAAGAGGCTCTCCCTT 1583

```

```

DB 845 CCGAGCCAAAGTGAACCTGTCTCCCAACTGAGTCTGACAAAGAGGCTCTCCCT 904
QY 1584 GATCAGCCCGCTCTGAAGAACAAGCCGCGCCCGGACGAGATGAGAGGAGGCG 1643
DB 905 GATCAGCCCGCTCTGAAGAACAAGCCGCGCCCGGACGAGAGCAAGAAAGAGGCG 964
QY 1644 GAGGAAGAGTTCCTCAAGTGAATAAGGCTTCAATGACTGAAGAAAGTCTACACGA 1703
DB 965 GCGAAAGAGATTCCTCAAGCAGCAAGGCTGATCTATGCTTAAGAAATCTCCACACGA 1024
QY 1704 GCGAATATCTGAAGAGATCTCGAAGTATCACTTGTGTTCAGACACAGTGAACA 1763
DB 1025 GCGCAGTACTGAAGAGCTTGAAGATCTGCTGTGTTCAGAGCAGTGAAGCA 1084
QY 1764 AGAGAGCCGATCCGGAAGCACTGAAGATCTCAATTCGGAATTTGAACCTTGA 1823
DB 1085 AGAGAGCTCATGCTCCGAAGCTTGAAGATCTCAATTCGGAATTTGAACCTTGA 1144
QY 1824 CAATTTCAATTAATTTCTCAAGAAATTGAGCAAGCACTTGCCTGTGGAGGCG 1883
DB 1145 CAAGTTCATACCAATTTCTGAAGAAATTGAGCAAGCACTTGCCTGTGGAGGCG 1204
QY 1884 CTCAAATGCCCAATGA--GAGATTAACAAGAAATCGGATGTCACTGTAAGAAAT 1940
DB 1205 CTCAAATGCCCAAGTCAAGAGAAATTACAAAGAAATCGGAGGATCATGCTGA 1264
QY 1941 TCAAGGCAATGAACACTGCGGCTCACTGTGGAAGCAAGAGAGCCCTTGAAGCCCT 2000
DB 1265 TCAAGGCAATGAACACTGCGGCTCACTGTGGAAGCAAGAGAGCCCTTGAAGCCCT 1324
QY 2001 GGAGAAATGAATCAAGAGCTCCGCGGCTGAGAACTTCTGAGAGACTTGAAGCTGA 2060
DB 1325 GGAGAAATGAATCAAGAGCTCCGCGGCTGAGAACTTCTGAGAGACTTGAAGCTGA 1384
QY 2061 GAAAGTGTGTAACCTACCTGCAACCTTCTCTGCGGCACTGCAAGGCTCATGA 2120
DB 1385 GAAAGTGTGTAACCTACCTGCAACCTTCTCTGCGGCACTGCAAGGCTCATGA 1444
QY 2121 CTACAGAGTCTTGAAGCGGCTGTGCAACCAAGCCCGGACGAGCCGAGCTTGA 2180
DB 1445 CTACAGAGTCTTGAAGCGGCTGTGCAACCAAGCCCGGACGAGCCGAGCTTGA 1504
QY 2181 GGAAGTGTGTAACCTACCTGCAACCTTCTCTGCGGCACTGCAAGGCTCATGA 2240
DB 1505 GGAAGTGTGTAACCTACCTGCAACCTTCTCTGCGGCACTGCAAGGCTCATGA 1564
QY 2241 GATCAAGATGAAGATTTCCAGAGCTGCAAGAACTCAAGAAAGATTTGAAGCTGA 2300
DB 1565 GATCAAGATGAAGATTTCCAGAGCTGCAAGAACTCAAGAAAGATTTGAAGCTGA 1624
QY 2301 CAATCTTGTGTTCCGGGAAAGGAGTTCAATCCGTCTGGGCAAGCTTCAAGAACT 2360
DB 1625 CAATCTTGTGTTCCGGGAAAGGAGTTCAATCCGTCTGGGCAAGCTTCAAGAACT 1684
QY 2361 GAAAGGCTTCAAGAGCTGATTTCTTCTGTTCAAGAGCTCTGTTATACAGAGCG 2420
DB 1685 GAAAGGCTTCAAGAGCTGATTTCTTCTGTTCAAGAGCTCTGTTATACAGAGCG 1744
QY 2421 GGGGCTGAGCGCTCAATCAATTAAGTTCAGGAGCTCCGCTCTATGAGCATGAC 2480
DB 1745 GGGGCTGAGCGCTCAATCAATTAAGTTCAGGAGCTCCGCTCTATGAGCATGAC 1804
QY 2481 GATTGAGAGAGGAAAGAGATGGGGGTGCCCCATGCTGACCTCCGGGGCCAGCG 2540
DB 1805 GATTGAGAGAGGAAAGAGATGGGGGTGCCCCATGCTGACCTCCGGGGCCAGCG 1864
QY 2541 GAGTCCATCATCTGTGGCGCAGTTCTGCTGCTGAGATGAGAAATGGGTTGAGACAT 2600
DB 1865 GAGTCCATCATCTGTGGCGCAGTTCTGCTGCTGAGATGAGAAATGGGTTGAGACAT 1924
QY 2601 CCAATGAGCATTTGACCTTGGCGAGAAAGACAGAGCCCGGCTGAGTTCTTGCAG 2660
DB 1925 CCAATGAGCATTTGACCTTGGCGAGAAAGACATGAGACCAAGCTTCTTGCAG 1984

```

Qy	2861	CAGGCCCCCTGACAAACAAGTCCCTGATGAAGACCAACGGGGCTGACCAAGAGTCAAGAGA	2728
Db	1985	CAGCCCCCTTACACAACAAGTCCCGGATGAAGCCACAGGGCTGACCAAGAGTCAAGAGA	2044
Qy	2721	TGACCTGAGGCCCTCTCGCGCACATGCTGTGAAGCGCCAGGCCCCCGCACCGCGCAACACAT	2780
Qy	2781	GGTGCAAGTGTGTGTGGCACCGGCAACCAAGGTCCTCCATGGTGAATTGACATTCGCAAT	2840
Db	2105	GGTCACAGTGTGTGTGGCATTCGACAGACCAAGTCTCCATGGTGAATTGACATTCGCTGT	2164
Qy	2841	GGAGAATCAGTTGTCTGGAAAACCTGCTCAGGAATTCAAAAACAGCAACGGGTGGCAGAA	2900
Db	2165	GGAGAATCAGTCTCTGGGAACCTGCTCAGGAATTCAGAAACAGCAATGGCTGGCAGAA	2224
Qy	2901	GCTGTGGGTGGTGTTCACAACTTCTGTCCTGTTCTTCTACAAATCACACGACAAATCA	2966
Db	2225	GCTGTGGGTGGTGTATTCACAACTTCTGTCCTGTTCTTCTACAAATCACACGACAAATCA	2284
Qy	2961	TCCCTTTGCCAGCTGCTCTGTGCTCTGCTACTCTGCTACCAATCCCTCTGAGTCCGAGAA	3020
Db	2285	CCCCCTTGCCAGCCCTGCTCTGTGCTGCTACTCTCTCACATCCCTCTGAGTCTGAGAA	2344
Qy	3021	CATCCAGAAAGACTACGTTCAAGCTSCACTTCAAGTCCGACGCTCTACTTCAAGGGC	3088
Db	2345	CATCCAGAAAGACTATGTGTTCCTCAAGCTSCATTCAAGTCCGACGCTCTACTTCAAGGGC	2404
Qy	3081	GGAAGACGAGTACACGTTCCGAAAGGTGATGGAAGTCCGACGCTGACCAAGCTCTGC	3140
Db	2405	TGAAGACGAAATACATTTGAAGAGGTGATGGAAGTCCGAAAGTCCGACCAAGCTCTGACG	2464
Qy	3141	CTCCGCAACCCCAACGTTGAGCCCAAGAAGTCTTGTGATTTGATGCGCGGACACACT	3200
Db	2465	CTCTCGGGCCCACTTGTGAGTCAAAAGAAATCTCACTGTACTGACGGCTGACACGCC	2524
Qy	3201	C 3201	
Db	2525	C 2525	
RESULT 8			
LOCUS	CQ724059	2937 bp	DNA linear PAT 03-FEB-2004
DEFINITION	Sequence 9993 from Patent WO02068579.		
ACCESSION	CQ724059		
VERSION	CQ724059.1	GI:42284916	
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;		
	Homnidae; Homo.		
REFERENCE	1		
AUTHORS	Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.		
TITLE	Kits, such as nucleic acid arrays, comprising a majority of		
	humanxons or transcripts, for detecting expression and other uses		
	thereof		
JOURNAL	Patent: WO 02068579-A 9993 06-SEP-2002;		
	PE Corporation (NY) (US)		
FEATURES	Location/Qualifiers		
source	1..2937		
	/organism="Homo sapiens"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:9606"		
ORIGIN			
	Query Match 53.8%; Score 1851.8; DB 2; Length 2937;		
	Best Local Similarity 79.9%; Pred. No. 0;		
	Matches 2513; Conservative 0; Mismatches 77; Indels 555; Gaps 6;		
Qy	220	CAMAGAGCTCTCGGAGAGGTGCTGCTGGATGCAAGTTTGCAACCACTCAACCTCGTGAA	279

Db	1	CAAAAGCTCCTGGGAAGGTGCTGCTGATGCAAGTTTGGCAACCACTCAACCTCTGTAAGAA	60
Qy	280	GGTAGCTATTTTGGGCTCGAGTTTCTGATCAGAAAAAGATCACGGTGTGGCTGGATCTC	339
Db	61	GGTAGCTATTTTGGGCTCGAGTTTCTGATCAGAAAAAGATCAAG-----	105
Qy	340	CTAAACCCTATGTGAAACAGATTTGAAGGCCAAAGACGTTGTTGATGTTGTGTG	359
Db	106	-----GCCAGGTTCCGGGTG	120
Qy	400	AAATTTCTTCGCGCTGACACACACATCTCAAGAAAGATCTCAAGAGTACCTGTGGC	459
Db	121	ACCTCTGCTGCTTTTGCTTGAGCGCCAGCTCTGTGATCTTGGCAGGTACCTGTGCGG	180
Qy	460	CTGCAGGTGAAGCAGAGCTTGGCTCAAGGACGTTGACGTGTATGTACACACAGCGACGT	519
Db	181	CTGCAGGTGAAGCAGAGCTTGGCTCAAGGACGTTGACGTGTATGTACACACAGCGACGT	240
Qy	520	CTCTTGATTTTCAACATTTGTGCATCTGAGATTTGGGGATTTTGTATGAAGCTTGGACAG	579
Db	241	CTCTTGATTTTCAACATTTGTGCATCTGAGATTTGGGGATTTTGTATGAAGCTTGGACAG	300
Qy	580	GAGCACTTAGCAAAAAATTAATACATACCTCAGCAGAACGCACTAGAGAGCAAAATCGT	639
Db	301	GAGCACTTAGCAAAAAATTAATACATACCTCAGCAGAACGCACTAGAGAGCAAAATCGT	360
Qy	640	GAATTTACCATACCAACATTTGACAAACACACAGCAGAAATCAGATTTCCAGCTCTTAG	699
Db	361	GAATTTACCATACCAACATTTGACAAACACACAGCAGAAATCAGATTTCCAGCTCTTAG	420
Qy	700	ATTGCGCGTGGGCTAGAGATGTATGAAATCCGGTTGACCCCGGCAAGGACAGGAAAGGC	759
Db	421	ATTGCGCGTGGGCTAGAGATGTATGAAATCCGGTTGACCCCGGCAAGGACAGGAAAGGC	480
Qy	760	ACGAAGATCAATCTGGCGCTTGCCACACGGGAATTTAGTGTTCAGGGTTTCACTAAG	819
Db	481	ACGAAGATCAATCTGGCGCTTGCCACACGGGAATTTAGTGTTCAGGGTTTCACTAAG	540
Qy	820	ATCAATGCTTTCAACTGGGCCAAGGTGCGGAAGCTGAGCTTCAAGAGAAAGCGTTTCTC	879
Db	541	ATCAATGCTTTCAACTGGGCCAAGGTGCGGAAGCTGAGCTTCAAGAGAAAGCGTTTCTC	600
Qy	880	ATCAAGCTCCGGGCCAGATGCGCATATGTCGATCAAGATACCTTGGAAATTCCTGATGGCC	939
Db	601	ATCAAGCTCCGGGCCAGATGCGCATATGTCGATCAAGATACCTTGGAAATTCCTGATGGCC	660
Qy	940	AGTGGGGATTTCTGCAAGTCCCTTGGAAAAATCTGTGTAAACATCAATGCTTCTTAGA	999
Db	661	AGTGGGGATTTCTGCAAGTCCCTTGGAAAAATCTGTGTAAACATCAATGCTTCTTAGA	720
Qy	1000	CTTTTGAAGAGCCCAAAACCAAAAGCCCAAGCCCGTCTCTTAGCCGGGGGTCAATCTTT	1059
Db	721	CTTTTGAAGAGCCCAAAACCAAAAGCCCAAGCCCGTCTCTTAGCCGGGGGTCAATCTTT	780
Qy	1060	CGGTTCAAGTGTGCGGATCTCAGAGCAGGTTCTGATCTATGTATAAGAAAGGACATTAAG	1119
Db	781	CGGTTCAAGTGTGCGGATCTCAGAGCAGGTTCTGATCTATGTATAAGAAAGGACATTAAG	840
Qy	1120	AAGGTGACGTTTGAAGAAAGACACAGCAAGATTCATCTATCCGGAGCTTGTCTTACAG	1179
Db	841	AAGGTGACGTTTGAAGAAAGACACAGCAAGATTCATCTATCCGGAGCTTGTCTTACAG	900
Qy	1180	CCTACAGAACTGAATTCGGAAGTCTGGAAGCAGTCTCAGCAGAGCACACAGCTTTACATTT	1239
Db	901	CCTACAGAACTGAATTCGGAAGTCTGGAAGCAGTCTCAGCAGAGCACACAGCTTTACATTT	960
Qy	1240	GGAGAAGGTGCGGATCTTCAAGGGGGCCAGAGCTGCGCGGAGGAAAGAAACGAAAGTT	1299
Db	961	GGAGAAGGTGCGGATCTTCAAGGGGGCCAGAGCTGCGCGGAGGAAAGAAACGAAAGTT	1020
Qy	1300	TCCGCCGGGGAGCCGGGGGTGCAACCCGAGGCTTGGCGGAGGAAGACCCCGGGGTAAAC	1359

VERSION BC004009.1 GI:13278387
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidae; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 2632)
AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schlier, G.D., Altchul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diachenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Scheet, T.E., Brownstein, M.J., Usdin, T.B., Tobiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loggellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Huiyk, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahy, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schultz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalls, D.E., Scherch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 2632)
AUTHORS Strausberg, R.
JOURNAL Direct Submission
PUBMED Submitted (28-FEB-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
TITLE NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
COMMENT Contact: MGC help desk
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-shgc.stanford.edu>
Contact: (Dickson, Mark) mcd@paxi.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
Series: IRAC Plate: 8 Row: P Column: 11
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis.
FEATURES
Location/Qualifiers
1..2632
/organism="Mus musculus"
/mol_type="mRNA"
/strain="Czech II"
/db_xref="taxon:10090"
/clone="IMAGE:3493093"
/tissue_type="Mammary tumor metastasized to lung. Tumor arose spontaneously from a senescent normal mammary (clonal) outgrowth infected with the virus MMTV."
/clone_1lb="NCI CGAP_Lu29"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
1..2632
/gene="Farpl"
/note="synonym: CdeP"
/db_xref="GeneID:223254"

CDS
/db_xref="MGI:2446173"
1..1263
/gene="Farpl"
/codon_start=1
/product="Farpl protein"
/protein_id="AAH04009.1"
/db_xref="GI:13278388"
/db_xref="GeneID:223254"
/db_xref="MGI:2446173"
/translation="YMLKNIQGMKLAHLWMSBALALENSIKSRLERCPDF LQVLYPLNTLPLRLHMYKIVLEKCHHPNADFPDCCRAALETBMAQL HQTMYKMFQLEHDKDLDIGDVLVGRFIRIGSLKSGKLGQRMPLFEND LLYTGRGLTASNOFVHGQLPLYGMTIESEBEWGPCHLTIRGROSITVAASNSE MEKMEMDIOMADLEAKNSGPPELLASPPDKSPDEATADQSEDDLSARTSE ROAHRGNTMVCHNRSTSVMPFSLAVENGLNLRKRNKSGMKLVVFTNF CLFPKSHODSRPLASLPILGYSLTIPRESENHIDYVKLHFKSHVYFRASEYTF ERKMEVINSATSSASRAHLHSHKESHLY"
55..309
/gene="Farpl"
/note="RhoGEF: Region: Guanine nucleotide exchange factor for Rho/Rac/Cdc42-like GTPases"
/db_xref="CCD:smart00325"
48..699
/gene="Farpl"
/note="PH: Region: Pleckstrin homology domain. Domain commonly found in eukaryotic signalling proteins. The domain family possesses multiple functions including the abilities to bind inositol phosphates, and various proteins. PH domains have been found to possess inserted domains (such as in PIV gamma, synutrophin) and to be inserted within other domains. Mutations in Brutons tyrosine kinase (Btk) within its PH domain cause X-linked agammaglobulinemia (XLA) in patients. Point mutations cluster into the positively charged end of the molecule around the predicted binding site for phosphatidylinositol lipids"
/db_xref="CCD:smart00233"
928..1218
/gene="Farpl"
/note="PH: Region: Pleckstrin homology domain. Domain commonly found in eukaryotic signalling proteins. The domain family possesses multiple functions including the abilities to bind inositol phosphates, and various proteins. PH domains have been found to possess inserted domains (such as in PIV gamma, synutrophin) and to be inserted within other domains. Mutations in Brutons tyrosine kinase (Btk) within its PH domain cause X-linked agammaglobulinemia (XLA) in patients. Point mutations cluster into the positively charged end of the molecule around the predicted binding site for phosphatidylinositol lipids"
/db_xref="CCD:smart00233"

ORIGIN
Query Match 30.6%; Score 1054; DB 6; Length 2632;
Best Local Similarity 89.0%; Pred. No. 3,5e-288;
Matches 1138; Conservative 0; Mismatches 140; Indels 0; Gaps 0;
1924 GTCATGCTGAAGAATCTTAGGGCATGAGCACTGGCGCTCACTGTGGAAGCACAGC 1983
1 GTCATGCTGAAGAAATCTTAGGGCATGAGCACTGGCGCTCACTGTGGAAGCACAGT 60
1984 GAGGCGCTTGAGAGCCCTTGAGAGATGGAATCAAGAGCTCCGCGCGCTGGAAGACTTCTGC 2043
61 GAGGCGCTTGAGAGCCCTTGAGAGACTTCATCAAGAGGCTCCGCGCGCTGGAAGACTTCTGC 120
2044 AGAGACTTTAGACTGAGAGAGGTGTGTTACTTACCGGCTCAACACCTTCTCTCGCGCCA 2103
121 CGAGACTTGAGACTGAGAGAGGTGTGTTACTTACCGGCTCAACACCTTCTCTCGCGCGC 180
2104 CTGACCGGCTCATGACTTACAGACAGGCTCTGAGAGCGGCTGTGCAAAACCCGCGC 2163
181 CTGACCGGCTCATGACTTACAGACAGGCTGTGAGAGCGGCTGTGCAAAACCCGCA 240

QY 2164 AGCCAGCCGCACTTCAGGGAATGCGGCGCTTTGGCAGAGATCAGGAGATGTGGCA 2223
 DB 241 AACCAAGCCCTCAGGAGCTGAGAGCTGCGCTGGCGGAGATCAGAGATGTGGCC 300
 QY 2224 CAGCTCCAGCGTACGATGATCAAGATGGAATTTTCAGAAAGCTGCAGAACTCAAGAAA 2283
 DB 301 CAGCTGACGGGACATGATCMAAGATGGAATCTTCAGAAAGCTGCAGAGCTCAAGAAA 360
 QY 2284 GATTGATTTGGCATGCAATCTTTGGTTCCGGGAAAGGAGTTCACTCGCTGGGAGC 2343
 DB 361 GATCTGATCGGCAATGCAATCTTTGATCCAGAAAGGAGTTCACTCGCTGGGAGC 420
 QY 2344 CTGAGCAAGCTTCGGGGAAAGGGGCTCCAGACGGCATTTCTTCCTGTTCAACAGCTC 2403
 DB 421 CTGAGCAAGCTTCGGGGAAAGGGGCTTCAGACGGCATTTCTTCCTGTTCAACAGCTC 480
 QY 2404 CTGCTATACAGAGCCGGGGCTGACGGCTCCCAATCAGTTTAAAGTCCAGGGCAGCTC 2463
 DB 481 TTGCTGATACAGCCGGGACTGACAGCATCTAATCAAGTTTAAAGTCCAGGGCAGCTC 540
 QY 2464 CCGCTCTATGAGCATGACGATTGAGAGAGCGAAAGCAGTGGGGGTGCCCACTGCTG 2523
 DB 541 CCACCTATGAGCATGACGATCGAGAGAGTGAAGAGAGTGGGGGTGCCCACTGCTG 600
 QY 2524 ACCCTCCGGGGCCAGCGGCGAGCTCATCATCTGTGGCCGCTTCGGGTCCGAGATGAG 2583
 DB 601 ACCCTTGAAGGGCCAGCGGCAATCTATCATGTGGCTCCGCTTCGCGTCAAGATGAG 660
 QY 2584 AAGTGGATTGAGAGCATCCAGATGAGGCACTTGAACCTGGCGGAGAGAGCAGACCCGCGC 2643
 DB 661 AATGATGAGAGAGCATCCAGATGAGGCACTTGAACCTGGCGGAGAGAGATGAGATC 720
 QY 2644 CTTGAGTTCTTGGCCAGCAGCCCCCTGACAAAGTCCCCTGATGAGGACACCGGCT 2703
 DB 721 CCGGAGCTCTTCCAGCAGCCCCCTGACAAAGTCCCCTGATGAGGACACCGGCT 780
 QY 2704 GACCAAGAGTCAGAGATGACCTGAGGGCTCCGCGCATCTGCTGAGAGCGCCAGCGCCG 2763
 DB 781 GACCAAGAGTCAGAGATGACCTGAGGGCTCCGCGCATCTGCTGAGAGCGCCAGCGCCG 840
 QY 2764 CACCGGCGCAACAATGATGAGCATGCTGTGGCAACCGCAACAACCGCTCTCATGATG 2823
 DB 841 CACCGGCGCAACAATGATGAGCATGCTGTGGCAACCGCTCTCATGATGATGATG 900
 QY 2824 GACTTCAAGATGAGATGAGATCAATGATGAGAACTGCTGAGAGAAATCAAAAAC 2883
 DB 901 GACTTCAAGATGAGATGAGATCAATGATGAGAACTGCTGAGAGAAATCAAAAAC 960
 QY 2884 AGCAAGGGTGGAGAAAGTGTGGTGTTCACAACTTCTGCTGTTCTTCAAAA 2943
 DB 961 AGCAAGGGTGGAGAAAGTGTGGTGTTCACAACTTCTGCTGTTCTTCAAAA 1020
 QY 2944 TCACACAGGACAAATGATCCCTTGGCAGGCTCTGCTCGGCTACTGCTCAGCATC 3003
 DB 1021 TCACACAGGACAAATGATCCCTTGGCAGGCTCTGCTCGGCTACTGCTCAGCATC 1080
 QY 3004 CCGCTGAGTCCGAGAAATCAGAAAGCTAGCTTCAAGTTCAGATCCAC 3063
 DB 1081 CCGTTCGAGTCCGAGAAATCAGAAAGCTAGCTTCAAGTTCAGATCCAC 1140
 QY 3064 GTTCTACTTTCAGGGCGGAAAGCGAGTACAGTTTGAAGAGTGAAGTGAATCCG 3123
 DB 1141 GTTCTACTTTCAGGGCGGAAAGCGAGTACAGTTTGAAGAGTGAAGTGAATCCG 1200
 QY 3124 AGTGCACAGCTCTGCTCGGACCCAGAGTGTGAGCCAAAGAGTCTTGTGTAT 3183
 DB 1201 AGTGCACAGCTCTGCTCGGACCCAGAGTGTGAGCCAAAGAGTCTTGTGTAT 1260
 QY 3184 TGATGCGGAGCACTC 3201
 DB 1261 TGACGGCTGAGCAGGCC 1278

RESULT 10
 CQ725044 3997 bp DNA linear PAT 03-FEB-2004
 LOCUS Sequence 10978 from Patent WO02068579.
 DEFINITION CQ725044
 ACCESION CQ725044
 VERSION CQ725044.1 GI:42285901
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homnidae; Homo.
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 Patent: WO 02068579-A 10978 06-SEP-2002;
 PB Corporation (NY) (US)
 FEATURES
 source
 Location/Qualifiers
 1..3997
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 ORIGIN
 Query Match 30.2%; Score 1039; DB 2; Length 3997;
 Best Local Similarity 60.3%; Pred. No. 7, 2e-284;
 Matches 1865; Conservative 0; Mismatches 1185; Indels 45; Gaps 7;
 QY 69 GCCGACCAGATCAGACTGGGGGCCCCGAAAAATTTGGGGATGATACCTTGGAAAG 128
 DB 150 GCAGACTGAGAGGATGCGTTGGGTGCCAGACCCCTGGGAGTTAGACCCCTGAGCC 209
 QY 129 TGGACGAAGCCGCCCAACACCTTCAGAAAACTCGTCCATCAATCAATCCAGATGCT 188
 DB 210 TGGGAGACTCTCTTCCAGATGAGAGAGAGACCTGCACTGAGTAAAGTGTGCT 269
 QY 189 GGATGACACCAGAGGACATTTGAAGTTCCAAAGAGCTCTGGAGAGTGTCTGTA 248
 DB 270 GGAACAACATGAGAAATTTGACATGAGCTTAAATGCAATGAGGCAAGTATTAATGAC 329
 QY 249 TGCATTTGCAACCACTCACTGATGAGTGTATTTGGCTGAGTTTCTTA 308
 DB 330 ACAAGTGTGAAGCGTTTAACTGTGATGATGATCTTCTGGATGAGATTTCAAAA 389
 QY 309 TCACAAAAAGATCAGCGTGTGGATCTCTTAAACCATTGTGAACAGATTGAAG 368
 DB 390 TACTCAGTCTTATGATTTGGCTTGAACCTTGAACCATTGATGAGCAATGAGAG 449
 QY 369 GCCAAAGCAGTTGTTGTTGAATTTGTGTGAATTTCTTCGCTGACACACAACT 428
 DB 450 GCCAAAGATGTGTGCTTCCGCTAGCTGTAATAATTTTTCACCTGATCTGTGAGCT 509
 QY 429 CCAGAAAGAACTCAGAGTACTGTTGCGCTGCAAGTGAAGCAGAGATTGGCTCAAG 488
 DB 510 ACAGAAAGAAATPACAGATCTTGTTCCTTGAACCTTGAAGAGACCTCTGAGAA 569
 QY 489 CAGGTGACGTGTAATGACACAGGAGCTCTTGAATTTGACATGTTGCAATCTGA 548
 DB 570 GCGTTTACCTGTGCTGACACACAGCGGCTCTTCAAGTCCATCTCTGAGTCGGA 629
 QY 549 GATTGGGATTTTATGAAAGCTTGGACAGAGACATTTAGCAAAAAATTAATACATAC 608
 DB 630 AATAGGAGATTAAGATGAAGACCTGGACCGAGAGACCTCAAGTGAAGATATTTGCC 689
 QY 609 TCAAGAAAGCAGACTAGAGAGCAAAATCGTGAATTTGACCAATCAATGAGCAAC 668
 DB 690 TGGCCAGCAGACTGCTTGAAGAAATTAATTAATTCATCAAGAGCAGTGGGCGAGC 749
 QY 669 ACCAGCAGATCAGATTTCCAGCTCTTGAAGATTCGCGCTGAGATGATGAT 728

Db 750 ACCTGCTGAGTGGGATTTCTCAGGTGCTCGAATAATGCTCGAATAATGATGACGCGAT 809
Qy 729 CCGGTTGCAACCGGCGCAAGAGACAGAGAGGACAGAGATCAATCTGCGCGTTGCCAACAC 788
Db 810 CAGATTTTCATGATGCTTCTGACAGGAGAGAGAACCAAGATTTCAACTGCGAGTTTCCACAT 869
Qy 789 GGGAAATTCAGTGTTCAGGGTTTCACTAAGATCAATGCTTCAACTGCGCGCAAGGTGCG 848
Db 870 GGGGTACTGATGTTTCAAGGACACCAAAATCAACATTTTCAACTGCTCCAGAGTCCG 929
Qy 849 GAAGCTGACTTTCAGAGAGAGGCTTCTTCATCAAGCTCCGCGCGAGTCCCAATGTGC 908
Db 930 TAAACTACCTTCAGAGAGAGAGAGATTTCTTACAACTTCACAGAGGTTCTATGACC 989
Qy 909 GTACAGAGATACCTTGGAAATTCCTGATGCGAGTCCGAGATTTCTGCAAGTCTCTGAA 968
Db 990 TTAGCGAGACACTTGAATTTTGTGGGTATGAGATGATGATGATGATGATGATGATGATGAT 1049
Qy 969 AATCTGTGTGAACATCATGCTTCTTTTGAAGCTTTTGAAGGCCCAACCAAGCCCA 1028
Db 1050 GATTTGTGAGATACACACCTTTTGTAGACTTTTGAACCACTTACCAAAAGCAAA 1109
Qy 1029 GCCCGTCTCTTTAGCGCGGGGTCATCATTTGCGTTCACTGCTGCGACCTCAGAGAGGT 1088
Db 1110 AGCGGCTTCTTCAAGCGGGGCTCTCTTCAAGATCAGTGAAGAACTCAGAAACACT 1169
Qy 1089 TCTCGACTATGTTTAAAGAGAGAGACATTAAGAGGTGACGTTTGAAGAGAGACAGCA 1148
Db 1170 AGTAGATTTTTCAGAGACAGTGAATGAGAGAAATTCATATGAGAGAGAGAGAGAGAG 1229
Qy 1149 GATTCATTTCTATCCGAGGCTTGTCTCAGAGCTTACAGAACTGAAATTCGAAAGTGTGA 1208
Db 1230 GACCCACAGTCC-----GTTGAGCTTGACTGACAGACCTACCAAA 1271
Qy 1209 GAGGTCTAGACAGAGACAGAGCTTACATTTGAGAGAGGTGCGAATCTCCAGGGGCGCA 1268
Db 1272 ACAGAGCATCTCATTTCCGAGGAGATTGAGACTCTGCTCTCCCACTTTCAGCGAATGC 1331
Qy 1269 GAGCTCCCGGCGAGAGAGAAACGAAAGTTTCCGCGGGAGACCGGGTCCGACCCGAG 1328
Db 1332 CTTTTCATGCTCTCTCCCTCCACTGATCCCCCTTGGCTGCGAGATTTTAAGAGACAG 1391
Qy 1329 CCCTGCGCGAGAGAGAGCCCGGGGTAAACAGAGCGAGCGAGCGGCTCGGCGCC 1388
Db 1392 CAGCAGCTCCCTCAGATCCCGAGGTTTCTTACGTAAGATCCAGCTGACAGAGAGCG 1451
Qy 1389 CAGCGAGAGAGAGAGAGAG---TGGTTAAGATAGAGACCCAGAGAGATTAACCTAGCC 1445
Db 1452 CAGTGAAGAGATGAG 1511
Qy 1446 CCGGAG 1505
Db 1512 ACTCCAGCTCTGATCCAG 1571
Qy 1506 CTGCGAG 1565
Db 1572 GAGCCCCCTGAGCTGAG 1622
Qy 1566 CAG 1625
Db 1623 ACAG 1682
Qy 1626 CGATGAGATGAG 1685
Db 1683 CTGCGAG---GAGCCAG 1739
Qy 1686 GGAAGTGTATACACAG 1745
Db 1740 AAGAGATTTCTGCTACAG 1799
Qy 1746 TCGAG 1805
Db 1800 CCGAG 1859

Qy 1806 GAATTTGAGACTTTGACAAATTTTATATTAATTTCTCAAGAGAAATGAGCAAGACT 1865
Db 1860 CAACATGATCCCATCTATGATTTTCAAGAGAGGCTTCTGCGAGAGTGAAGCAAGGCT 1919
Qy 1866 TGCCTGTGGAGAGAGCGCTCAATATGCCAATCAGAG---ATTTCAGAAAGATGGCGA 1922
Db 1920 GGCACCTTGGAG 1979
Qy 1923 TGTATGCTGAAGAGATTTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1982
Db 1980 CATCTGCTCAG 2039
Qy 1983 CGAGGCTTGGAG 2042
Db 2040 CGAGGCTTCAAG 2099
Qy 2043 CAGAGACTTTGAG 2102
Db 2100 CAG 2159
Qy 2103 ACTGACCGGCTCATGCACTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2162
Db 2160 CATCAG 2219
Qy 2163 GAGCAGCGCGACTTCAAG 2222
Db 2220 CGGAGACATGACTACGCTGAG 2279
Qy 2223 ACAGCTCCAG 2282
Db 2280 CACACTTACAG 2339
Qy 2283 AGATTTGATGAG 2342
Db 2340 GAGCTGTGAG 2399
Qy 2343 CCTCAG 2402
Db 2400 CTTTCAAG 2459
Qy 2403 CCGTGTATACAG 2462
Db 2460 GTTGTGTACAG 2519
Qy 2463 CCGGCTATGAG 2522
Db 2520 TCCCTTCAAG 2579
Qy 2523 GACCTTCCGAG 2582
Db 2580 CACCATTAAG 2639
Qy 2583 GAGTGTGATGAG 2642
Db 2640 GAAATGATGCTGAG 2699
Qy 2643 CCTGAGTTCTGAG 2696
Db 2700 CCTGCACTGAG 2759
Qy 2697 CGGAGCTGACAG 2756
Db 2760 TCTG---GAGCAG 2816
Qy 2757 GGCCTCCGAG 2816
Db 2817 TGGCAG 2876
Qy 2817 CATGTGAGACTTCAAG 2876
Db 2877 CAGGAG 2936

QY	2877	AAAAAACAAGAACGGGATGGCAAGAACTGGGGTGGTGTTCACAAACTTCGCTGTTCTT	2936
Db	2937	CAAAAACAGTCATGGCTGGCAGAAAGCTTGGGATGCTCTTTACCAACTTCTGTGTGTTCTT	2996
QY	2937	CTACAAATCACACACAGCACAATTCATCCCTTTGCGACGCTGCTGCTCGGCTACTCGCT	2996
Db	2997	CTACAAAACTCATCAGATGATGATCACCTGGCCAGCTCTCCGGCTGCTGGGCTACAGGCT	3056
QY	2997	CACCATCCCCCTCTGAGTCCGAGAACATCCAGAAAGATTAAGTGTTCAAAGCTGCACCTTAA	3056
Db	3057	GAGCATCCCCCAGGGAGGCGCATGGCATACAAAGACATATTTTTCAAAGCTCCAGTTCAA	3116
QY	3057	GTCCCAAGCTACTACTACTTCAGGGCGGAAGCGAGTACACTTTCGAAAGAGTGATGGAAGT	3116
Db	3117	ATCCCAAGTCACTTCTTCCGGGGCTGAGAGAGCAATTTGAAAGGTGGATGGAGGT	3176
QY	3117	GATCCGCAAGTCCACACAGGCTCTGCTTCGCAACCC	3151
Db	3177	GATCCAGGGGCGCAGCAGCTCAGCCGGAGGGGCC	3211

[illegible]

REFERENCE
1
AUTHORS Certà, U., Foser, S. and Meyer, K.
TITLE Transcriptional activity assay
Patent: EP 1471153-A 16 27-OCT-2004;
F. HOFFMANN-LA ROCHE AG (CH)

FEATURES	Location/Qualifiers
source	1. .3997

ORIGIN

Query Match	30.1%	Score 1037.4;	Length 3997;
Best Local Similarity	60.2%	Pred. No. 2.1e-283;	
Matches 1864; Conservative	0;	Mismatches 1186;	Indels 45; Gaps 7;

QY	69	GCACACCCCGAGATCACAGACTGGGGGGCCCCGGAAAAATCGGGGATTCAGTACTTGGAAAG	128
Db	150	GCACACTGACAGGAGATGCGCTTGGGGGCCCAACCCCTGTGGAGTATACACCCCTTGAACC	209
QY	129	TGACACAGAAAGCGCCCCCAACACCTTCAGAAAACTCGTGTCCATCAAAATCCAGATGCT	188
Db	210	TGGCACAAGCTCTTGTGCCAAGATGACAAGAAAGACCTGCACCTCAGAGTAAAGCTGCT	269
QY	189	GGATGACACCCAGGAGGCAATTGAAAGTCCACAAAGAGCTCTGTGGAAAGGTGCTGCGA	248
Db	270	GGACACACCATGAGAAATTTTGCATTGAGCTTAATGCGATGGCCAGATTAATCTGAC	329
QY	249	TGCAGTTTGACCAACCACTCAACCTCGTGGAAAGTGACTAATTTTGAAGCTCGAGTTTCCCTGA	308
Db	330	ACAAAGTGGAAAGGTTTAAACCTGGTGAATGTGACTACTTCGGGATGGAGTTTCAAA	389
QY	309	TCACAAAAGATCAAGGTGTGAGTGGATCTCTCAAAACCATTTGGAACAGATTAGAG	368
Db	390	TATCTAGTCCATCTGAGATTTTGGCTTGAACCTATGAAACCATTCATTAGGCAAAATGAGAG	449
QY	369	GCCAAAGCAGTTGTTTAAAGTTTGTGTGAAATCTTTTCGCGCTGACCAACAACAAT	428
Db	450	GCCAAAGAAATGTGTGCTTCGCTCAGCTGATGAAATTTTTCACACTGATCTCTGTGACCT	509

OY	429	CCAAAGAAAGACTCAAGAGTCTCTGTTCCGCTGCAGGTGAAGAGACTTGGCTCAAG	488
Db	510	ACAAAGAAATATCAAGATATCTTTGCTCTTCAACTTAAGAGAGCTGCTGAAGA	569
OY	489	CAGGTGACGTATATGACACACAGCGAGCTCTCTTGATTTACACATTTGTCATCTGA	548
Db	570	GCCTTTGACCTGTCTGACACCAACAGCGGCCCTTCTCACGTTCCATCTCTGCAGTGGG	629
OY	549	GATTGGGATTTTGATGAGCGCTTGGACAGAGACACTTGAACMAAAATAATATCATACC	608
Db	630	AATAGGAGATTACATGTAACGCTGGACCGAGAGACCTCMAAGTGAACGATATTTGGCC	689
OY	609	TCAGCAGAAGCGCACTAGAGACAAATACTGTGAATTTCAACATACCAATTTGACAAAC	668
Db	690	TGGCCAGACGACCTGCTTGAAGATATCTGAATTCATCAGTAAGAGACGTTGGGCCAGAC	749
OY	669	ACCGAGCAATCAGATTTTCCAGCTCTTGAAGATTGCCCGTGGCTAGAGATGTATGAAT	728
Db	750	ACCTGCTGAGTCGGAATTTCCAGGTGCTCGAATTTGCTGAAGTTGGAATGTACGGCAT	809
OY	729	CCGGTTGACCCCGGCCAAGAGCAGGGAAGGCAACAGATCAATCTGSCCTTGCCCAAC	788
Db	810	CAGATTTCAATGCTCTTGAACGAGGAAGGAACCAATTTCACTGGCAGTTTCCCAT	869
OY	789	GGGAATTCATGTTTCAAGGTTTCACTAAGATCATGCTTCACTGGGCCAAGSTGCG	848
Db	870	GGGTGTACTCGTTTCCAGGGCACCAACAAATCACACTTTCACGTGTCCAAAGTCCG	929
OY	849	GAACTGAGCTTCAAGAGGAAGCGCTTTCTCATTAACCTCCGGCAGATGCCAATAGTC	908
Db	930	TAAACTAAGCTTCAAGAGGAAGAAATTTCTTATCAACTTCATCAGAGGTTCAATGACC	989
OY	909	GTACCAAGATTCCTTGAATTTCTGATGAGCCAGTCGGAATTTTGCAGATCTTCTGAA	968
Db	990	TTACCAAGACACATTAAGATTTTGTGGGTAGTAGAGATGAATGTAAAGAACTTCTGAA	1049
OY	969	AATCTGTGTGAACATCATGCTTCTTTAAGCTTTTGAAGGCCAAACCAAGCCCA	1028
Db	1050	GATTTGTGTGAGATCACACTTTTGTAGCTTTTGAGCCCACTAAGCCCAAGCAAA	1109
OY	1029	GCCGCTCTCTTTAGCCGGGGGTCACTATTGCTTGAAGTGTCCGACTAGAGAGCAGGT	1088
Db	1110	AGCGCTCTTCTTCAAGCGGGCTCTCTCTTCAAGTACAGTGAAGAACTCAAGAAACA	1169
OY	1089	TCTCGACTATGTTAAAGAGAGGACATTAAGAGGTSCAGTTTGAAGAGACACACAA	1148
Db	1170	AGTAGATTTATTTCAAGACAGTGAATGMAAGAAATTCATATGMAAGAAAGGACACGA	1229
OY	1149	GATTCATTTCTATCCGAGAGCTTGCTTCAACGCTTACAGAACTGAATTTGGAAGTGTGA	1208
Db	1230	GACCCACACGTC-----GTTTCGAGCTCTGACTGACGACCTTACCA	1271
OY	1209	GCACTTTCAGAGAGCACACAGCTTATCATTTTGAAGAAAGTGCAGATCTTCAAGGGGCCA	1268
Db	1272	ACAGAGCATCTCATTTCCCGAGGGATGAGAGACTCTGCTTCCCATTTTCAAGGAATGC	1331
OY	1269	GAGCTGCCCGGAGAGAAAGAAACGAAGTTTTCGCGCGGGAAGCGGAGTTCGACCCGAG	1328
Db	1332	CTTTTACTCTGCTCTCTCCCTTCACTCTGTGTCCCTCTGTGCTGCAAGATTTAAGACAG	1391
OY	1329	CCCTGCGCCGAGAGGAAGCCCGCGGAGTAAACAGACAGCGAGCGAGCGCTTCGAGCGC	1388
Db	1392	CAGCAGCTCCCTCACAATCCCGAGGTTCCTAAGTCAAGATCTCAGCTGACAGAGAGCG	1451
OY	1389	CACGAGAGAAAGAGAGAG---TCGTTTAAAGATAGAACCCAGCAGAGTAACTTCAGCC	1445
Db	1452	CAGTGAAGCAGTGTCTGAGAGCCCGCACACCAATCGGACCAAGCCCTTCGAGGCCCGCG	1511
OY	1446	CCGCGACAGCAAGCAGAGTCTCCCTGTAATGGAGTCCCTCACTTTCCGAGCTGTCTGAA	1505
Db	1512	ACTTCAAGCTGTGTCAAGGCCCTTTTCAAGAAAGTCTTCAAGCTTCTTCCCTTCAGCGGAA	1571
OY	1506	CTCCACAGGGAGATGGCCCTCTGCAACGTAGCTTGTCTTCCAACTGAGCCCGAC	1565

```

Db      1572 GAGCCCCCTGAGCTGAGCCCTGATTTGAGGTCCTT-----TGAGCCAGCTGA 1622
Qy      1566 CAGCAGGCTCTCCCTTATGAGCCCCGCTGTAATGACAGAGCTGCCCCCGACGGA 1625
Db      1623 ACAGGGCTATCCCACTCTGAGCCCTCTCTCATGTATGCTGGGGAGCCGGATGGA 1682
Qy      1626 CGATGAGATGAGGGCCGAGAGAGATTTCCCACTGAATAAGCTTACTTATAGCTTA 1685
Db      1683 CTGCGAG---GAGCCAGACAGAGCCGCTGTGAGAGAGAGCTTACTTATAGTCAA 1739
Qy      1686 GGAAGTGTATACACCGAGCGAATATATCTGAAGATCTGAAGTTATCACTTCGTGTT 1745
Db      1740 AAGAGATTCCTGCTACAGAAAGAACTACCTCAAGAGATTAGAAGTATATACGTGTGTT 1799
Qy      1746 TCAGAGCAGAGTGAAGAGAAAGAGAGCCCATGCGGAGAGACTGMAAAAGTCTATATCC 1805
Db      1800 CGCAGCGAGAGTGTGAAGAGAGAGCGCATGCTGCACTGTAGTACGCTGCTCTTCTC 1859
Qy      1806 GAATTTTGAACCTTTGACAAATTTTATTAATTTTCTCAAGAAATTTGACAGCAACT 1865
Db      1860 CAACATCGATCCCATCTATGAGTTCCACAGAGGCTTCTGCGGAGGTGAGAGAGGCT 1919
Qy      1866 TGCCTGTGGAAAGGCCGCTCAATGCCCAATCAGAG---ATTACAAAGAAATGCGCA 1922
Db      1920 GGCACCTGTGGAAAGGCCCTCCAAAGCCCAACAAAGAGCATCATCAAGATCGGGGA 1979
Qy      1923 TGTGATGCTGAAGAACTTCAAGGCGATGAAGCCTGGGGGCTCAGCTGTGAAGCAG 1982
Db      1980 CATCTGCTCAGAAACATGCGCCAGTTAAAGAGTTTACAGCTACTTCCAAAGACATGA 2039
Qy      1983 CGAGGCTTGAAGGCCCTGAGAGATGAAATCAAGAGCTCCGGCGCTGAGAGAACTTCTG 2042
Db      2040 CGAGGCTCTTAACGAATCTGAAAAGGCTACAAAGCTGTAAAGAAATGAGGAGGTGA 2099
Qy      2043 CAGAGCTTTAGCTGCAAGAGTGTGTTAATCACTGCTCAACACTTCTCTGCGGCC 2102
Db      2100 CAAGGATTTGAGCTGCAGAAAGGTCTGCTAATTCCTCAACACTTCTCTGAGAGCC 2159
Qy      2103 ACTGCACTGGGCTCATGCACTAACAAGAGGTCCTGAGGGGCTGTGMAAACACACCGGCC 2162
Db      2160 CATCCAGCGGCTGCTGACACTACCGCTGCTGCTGCGCCGCTATGCGACATTTACGCC 2219
Qy      2163 GAGCCACGCCGACTTTCAGGAGACTCCGAGCCGCTTTGGCAGAGATCAGAGATGTGTC 2222
Db      2220 CGGGGACCATGACTAGCTGAGCTGACAGCGCCCTGAAGCATCAGAGGTGACAC 2279
Qy      2223 ACAGCTCCACGGTACGATGATCAAGTGAAGATTTTCCAGAGCTGCAGAACTCAAGA 2282
Db      2280 CACACTACAGCACTTCTCATCCGCTGAGAACTGCAAGAGCTTACGAGGTGACAGCG 2339
Qy      2283 AGATTTGATTGGCATTGACATCTTGTGTTCCGGGAAGGAGTTCATCGTCTGGGCGAG 2342
Db      2340 GAGCTGTGGGATGAGAACTCATTTGCTCTGCGAGAGAGTTCATCCGAGGGCTG 2399
Qy      2343 CCTCAGCAAGCTCTCGGGGAAAGGGGCTCCAGCAGCGCATGTTCTTCTGTTTCAAGCAG 2402
Db      2400 CCTTCACAGCTCACCAAGAAAGGCTGAGCAGAGAGATGTTTTTCTGTTCTCAGATAT 2459
Qy      2403 CCTGCTATACAGAGCCGGGGGCTGACGGCTTCATAGTTTAAAGTCCACGGGAGCT 2462
Db      2460 GTTGTGTCACAGAGAAAGAGTTCAGAGGACACACCTTCGGATCCGGGGCTCTCT 2519
Qy      2463 CCGGCTCTATGGCATGACATTTGAGAGAGAGCAAGAGTGGGGGCTGCCCACTGCTC 2522
Db      2520 TCCCTCTCAAGGAGTCTGTGGAAGAAAGTAAACGAGTGTCTGTTCCACACTGTTT 2579
Qy      2523 GACCTCTCGGGGCGACGCGCATCTCATCTGTGCGCCGCACTTCTCGGTCCGAGATGA 2582
Db      2580 CACCATCTACGGGGCTCAGAAAACATCTGAGTGGCAGCAGACCTCGGTGAGAAAGA 2639
Qy      2583 GAAGTGGGTTGAGGACATCCAGATGGCCATTTGACTGGCGGAGAGAGACAGACCCCGC 2642

```

```

Db      2640 GAAGTGAATGCTGAGCTTGAATCCGCGCATCCAGCAGCAGAGTGGCGGTGACACGGC 2699
Qy      2643 CCTGAGTTCCCTGGCAGCA-----GCCCCCTGAGCAACAAGTCCCTGATGAAGCCAC 2696
Db      2700 CCTGACATGCGAGGCCGCACTGTGTGCACTGTCTCCCCAGATCCCCCAAGAGATATC 2759
Qy      2697 CGCGGCTGACAGAGTGAAGATGATGATGAGCCCTCGGCGACATGCTGAGAGCGCA 2756
Db      2760 TCTG---GAGCAGAGATGAGAAAGATGATGCTCGGGGTGTCCGACAGCTCCCTGAGAGGCA 2816
Qy      2757 GGGCCCGCACTGGCGGCAACAATGTGTGCAAGTGTGTCTGCGACACCGCAACACAGCTCTC 2816
Db      2817 TGGCAGAGACCGGGCCAAACACCAATGACAGTGTGTGTAACCGGAACACAGCGTGTG 2876
Qy      2817 CATGTGAGACTTACAGCATCGCAGTGAAGATCAGTTGTCTGGAACCTGCTAGAGAAAT 2876
Db      2877 CAGGAGACACACAGTGAAGCTGTGAGAAACAGCTTTCAAGATATCTGTGAAGAAAT 2936
Qy      2877 CAAAAACAGCAACGGGTGGCAGAAAGCTGTGGGTGTTCACAACTTCTGCTGTTCTT 2936
Db      2937 CAAAAACAGTATGCTGGCAGAAAGCTGTGGGTGTCTTACAACTTCTGTTGTTCTT 2996
Qy      2937 CTACAAATCAACACAGACATCAATCCCTTCCAGACCTGCTGCTGCTATCCCT 2996
Db      2997 CTACAAACTATCATGAGATGACTACCACTGACGCTCCGCTGTGGGCTACAGGCT 3056
Qy      2997 CACATCCCTCTGAGTCCGAGAACATCAGAAAGACTAGTGTCAAGCTGCACCTCA 3056
Db      3057 GAGCATCCCCCAGAGAGGCGCATGTGATACAAAGACATATGTTTCAAGCTCCACTTCA 3116
Qy      3057 GTCCACGCTTACTACTTTCAGGGCGGAAAGCAGATGACGTTTGAAGGTGATGAAGT 3116
Db      3117 ATCCACGCTTACTTCTTCCGGGCTGAGAGCAAGTACAAATTTGAAGGTGATGAGGT 3176
Qy      3117 GATCCGAGTGCACCAAGCTCTGCTGCGACCC 3151
Db      3177 GATCCAGGGGCGCAGACGCTCAGCCGGAGGAGCC 3211

```

RESULT 12

AB018336 3997 bp mRNA linear PRI 10-JAN-2004

LOCUS AB018336 Homo sapiens mRNA for KIAA0793 protein, partial cds.
 DEFINITION
 AB018336
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homidae; Homo.

REFERENCE
 1 Nagase, T., Ishikawa, K., Suyama, M., Kikuno, R., Miyajima, N.,
 Tanaka, A., Kotani, H., Nomura, N. and Ohara, O.
 Prediction of the coding sequences of unidentified human genes. XI.
 The complete sequences of 100 new cDNA clones from brain which code
 for large proteins in vitro
 JOURNAL
 DNA Res. 5 (5), 277-286 (1998)

PUBMED 9872452
 JOURNAL
 2 (bases 1 to 3997)
 Ohara, O., Suyama, M., Nagase, T., Ishikawa, K. and Kikuno, R.
 Direct Submission
 Submitted (08-OCT-1998) Osamu Ohara, Kazusa DNA Research Institute,
 Laboratory of DNA Technology, Yana 1532-3, Kisarazu, Chiba
 292-0812, Japan (E-mail: cdna@fokkazusa.or.jp, Tel: +81-438-52-3913,
 Fax: +81-438-52-3914)

REFERENCE
 AUTHORS
 TITLE
 JOURNAL

FEATURES
 source
 1. 3997
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="hk05692"
 /sex="male"

gene
/tissue type="brain"
/clone lib="pBluescriptII SK plus"
/dev stage="adult"
1.3997
/gene="K1A0793"
<115.3382
/gene="K1A0793"
/note="Start codon is not identified"
/codon_start=1
/product="K1A0793 protein"
/protein_id="BA034513.2"
/db_xref="GI:40788367"
/translation="RMGEIEGTAVLQTAGRLAQAQTPVGVSTLEPQGLTPRMQEKH
LHRLVLDLMTMEIPDIEPKCDQVLLTQWKKLIVECQPMQENQTSVYMLTEP
MKETIQRIRPKVIVRIAKPPRPDQGOQEBYTRYLPQLOKRLLERLCALDTT
MALTSHLQSEIGDYDETLDREHLKNEYLPQGHLEKLELPHQHQVQALAESEDF
QVLEIARKLEMYGIRFPMASDRGTQIKQLAVSHMGVLFVGGTQIKITPMNSKVRKLSF
KRRFLIKLAEVHGVQDTLELGLSRECKRPMKICVEYHFFPLDLPKPKARAV
FPERGSFRRSGTQKQVDFPDQSGMKRIPYRRHSKTHSVTLADLPKOSI8FP
EGRLTPASPSANAFYLSPSLTVPSGLPFRKDSLSLNDPOVSVKSPAPBRSGAV
AGSPDRSAQPLQPALQPGRLSPGSLSPSSRSKSLSLSPASPOVPIBPAQSGSP
LSPVUSDAGAGMDCEBRPKRVPADPAEATFYKEILATERITLKLELITVPRSAV
VEDAMPATLMTLFSNIDPIYEFHNGFLKEVEQRLALWEGPSKATKSHQI8GL
LNRMRQLKEFTSYFORHDEVLELEKATKCKLEAVYKEFELQKCYLPANTFLKP
IORLHYRLRLRCGYSPGHHDYADCHDALKAITEVTTLOHILRIENLOKLTLE
ORLVGIBNLIARGREIRREGCHKLTKKGLQORMFPLSDMLLYTSKVGAGSHRI
RGLRPIQGMIVESDNEMSVPHCTTYIAAKTIVVASTRLEKMKMLDNL8IQAK
SGSDIAPALPGRVCTRPSPSPNEVSLQESBDPAQVNSLSLGHQOHRAHTTWHQW
YRNTSVSRADHSAVENQSLGVLKRFKSHQWKLWVFNTCLFPEYKTHODYPLA
SLPLGSVSI8PREADGINDYVFKLFQKSHVFPFAESKYTERMMEVYIQA8SSSAG
RAPSIVQDGPQSSGLEGMVRKEE"

ORIGIN

Query Match 30.1%; Score 1037.4; DB 5; Length 3997;
Best Local Similarity 60.2%; Pred. No. 2; Le-283;
Matches 1864; Conservative 0; Mismatches 1186; Indels 45; Gaps 7;
QY GCCGACCCCGAGATCACTGAGGCGCCCGGAAATTTGGGGATCGATCTTTGGAAG 128
DB TGCGACGTCGAGGAGTGGCTTGGGTGCCAGACCCCTGTGGAGTTTGACCCCTTGCC 209
QY TGCGACGAGAACCCGCCCAACCTTCAGGAAATCTGTGTCATCAAAATCCAGATGCT 188
DB TGCGACGACTCTCTTGCCAGAAATGCAAGAAAGCACTGCACTCAAGATTAAGCTGCT 269
QY GGAATGACCCAGAGGCACTTTGAAGTTCAAAAGAGCTCTGGGAAGGTGCTGCTGCA 248
DB GGAACACACCATGGAATATTTGACATTTGAGCCTTAATGCGATGGCCAGGTATTAAGTAC 329
QY TGCGATTTGCAACACCTCAACCTCGTGAAGGTAATTTTGGCTCGAGTTTCTGTA 308
DB ACAAGTGTGAAGCGTTTAAACCTGGAGATGAGTACTCTGGGAGTGAAGTTTCAAAA 389
QY TCACAAAGAGTCAACGCTGTGCTGTGATCTCTAAACCCATTGTGTAACAGATTAGAG 368
DB TACTCATGCTCACTGAGTTTGGCTTGAACCTAAGAACCATTAAGCAAAATAGGAAG 449
QY GCCAAAGCAGCTGTGTTTAAAGTTGTGTGAATTTCTTTCGGCTGACCAACAACA 428
DB GCCAAAGAAATGTGTGCTGCTAGCTGTAATTTTTCACACTGATCTGTGCTGACT 509
QY CCAAGAGAACTCAACAGTAACTGTTGCGGCTGAGGTCAGGTAAGCAGAGACTTGCTCAAG 488
DB ACAAGAAAGAAATTAACAGATATCTTGTGCTTGAACCTTAAGAGAGACTGTGGAAGA 569
QY CAGGTGACCTGTATGACACAGCGCAGCTCTTGTGATTTCAACATTTGTGCAATCTGA 548
DB GCGTTTGACCTGTGCTGACACAGCGGCTTCTCAAGTCCCATCTCTGCAAGTCCGA 629
QY GATTGGGATTTTGTATGAGCTTTGACAGAGGCACTTACGCAAAATTAATTAATCAATACC 608
DB AATAGGAGATTAGATGAACGCTGACCGAGAGCACTCAAAAGTGAACGAGATTTTGGCC 689

QY TCACAAAGCAGCACTAGAGAGCAAAATCGTGAATTTCAACATTAACATTTGAGCAAC 668
DB TGCGACGACGACTGCTTGAAGAAAGTACTAGAAATTCATCAGAACGATGGGCCAGAC 749
QY ACCAGCAGATCAGATTTTCAAGCTCTTAGAGATTTGCGCGCTAGAGATGTATGAAT 728
DB ACTGCTGAGTGCAGATTTTCAAGGTGCTCGAAATTTGCTCGAAAGTTGGAATGTAACGGCAT 809
QY CCGGTTGCAACCGGCCAAGACAGGGAAGGACAGAGATCAATCTGCGCGTGGCCAAAC 788
DB CAAATTTCAATGAGCTTGTGACAGGGAAGAACCAAGATTTCAACTGACGTTTCCACAT 869
QY GGAATTTCAATGATTTTCAAGGTTTCACTAAGATCAATGCTTCAACTGAGGCCAAGTGG 848
DB GGGTGTACTGTGTTTCAAGGACACCAAAATCAACATTTCAACTGTGTTCCAAAGTCCG 929
QY GAACTGACCTTCAAGAGAAAGCGCTTTCATCAAGCTCCGGCAGATGCCAATAGTGC 908
DB TAAACTAAGCTTCAAGAGAAAGAAATTTCTTATCAAACTTCAATCCAGAGTTTCATGAGCC 989
QY GTTACAGGATACCTTGGAAATTCCTGATGGCCAGTGGGATTTCTGCAATCTTCTGAA 968
DB TTAACAGGACACATTAAGAAATTTTGTGGGTGTAGAGATGAATGTAAAGACTTCTGAA 1049
QY AATCTGTGTTGAACATCATGCTTCTTGAACCTTTTGAAGAGCCCAACCAAGCCCAA 1028
DB GATTGTGTGAGATATCAACCTTTTGAACCTTTTGAACCACTTAAGCCAAAGCCAA 1109
QY GCCCGTCTCTTTAGCGCGGGGCTCATCTTTCGTTCACTGCTGAGTCCAGAGCAGGT 1088
DB AGCGGCTTCTTCAAGCGGGGCTCTCTTCAAGATCACTGGAAGAACTCAGAAACAAT 1169
QY TCTGCACTATGTTAAAGAAAGAGACATTAAGAGGTGCACTTTGAAGAGACAGACGA 1148
DB AGTAGATTTATTTCAAGAGAGTGAAGAAATTCATATGAAGAAAGGACACGCA 1229
QY GATTCAATTCATCCGAGGCTTGTCTTCAAGCTTACAGAACTGGAATTCGAAAGTCTGGA 1208
DB GACCCACAGTGC-----GTTCAGACTGTGATGTCAGACCTTACCAAA 1271
QY GAGTCTCAGACAGACGACGCTTATCAATTTGGAAGAGGTGCGAATCTTCAAGGGGCGCA 1268
DB ACAGACATCTCATTCGCCAGAGGATTTGAGACTCTGCTGCCCTCCATCTTCAAGCGAATGC 1331
QY GAGCTCCCGGCGAAGAAAGAACCGAAGTTTCCGCGGGAGCCGGGCTGCCACCCGAG 1328
DB CTTTTACTGCTCTCTCTCCCTCACTCTGCTCCCTTGTGCTGCGAAGTTTAAAGACAG 1391
QY CCTTGCGCGAGAGAAAGCCCGGGGTAAACAGCAGGCGGACGAGCGCTCGGCGGC 1388
DB CAGAGCTCTCTCAGATCTCCAGGTTTCTTAAGTCAAGAGATCCAGCTGCAAGAGAGGCG 1451
QY CACGAGAAAGAGAGAGAGG---TGTGTTAAGATAGAACCCAGCAGAGTAACTTCAGGCC 1445
DB CAGTGAAGAGATGCTGAGAGCCCGACACACATGAGGCCAGCCCTCGGGCCGCCGAC 1511
QY CCGGAGCCAGACAGAGCTCCCTGACTGAGCTGCTCACTTTCCAGACTGTCTGTGA 1505
DB ACTTCAGCTGTGCTCAGGCGCTTTCAGAGAGTCTTCAAGCTTCTCTTCCAGCCGGA 1571
QY CTTCGAGGGGAGAGTGGCCCTGCAACGTTGCTTCCCAACTGAGCCCGACAC 1565
DB GAGCCCTGAGTCTGAGCTTCAATTTCAAGTGCCTT-----TGGGCCAGCTGA 1622
QY CAAGCAGGCTCTCTCTTATGACGCGCTGTGTAATGACAGGCTGCCCCGAGCGA 1625
DB ACAGGCTCATCCCACTCTGAGCCCTGCTCTCAAGTATGCTGAGGAGAGCGGGATGGA 1682
QY CGATGAGATGAGGCGCGGAGAAAGATTTCCCACTGATTAAGGTATCTTCAATGCTTA 1685
DB CTGCGAG---GAGCCAGACACAGGCGCTGCTGAGCGAGGCTTATCTTAATGCTCAA 1739
QY GGAAGTGTACACACGAGCAGAAATATCTGAAGATCTCGAAGTTATCACTTCGTGTT 1745

Db 1740 AGAGATTCGTGACAGAAAGAAATACCTCAAGATTTAGAAAGTTATATACCGTGTGGTT 1799
 Qy 1746 TCAGACACAGTGTAGCAAAAGAGAGCCCATGCCGGAAGCACTGAAAAGTCTATATCCC 1805
 Db 1800 CCGCAGCGAGTGTGTAGAGAGAGCCCATGTCTGGACTGTGATACCTGTCTTTCTC 1859
 Qy 1806 GAATTTTGAACCTTGTGACAAATTTTATCTAATTTTCTCAAGAAATTGAGCAAGACT 1865
 Db 1860 CAACATCGATCCCATCTATGATGTTCCACAGAGGCTTCTGCGAGAGTGAAGAGAGGCT 1919
 Qy 1866 TGCCCTGTGGAAAGCCGCTCAAAATGCCCAATCAGAG--ATTACAAAGAAATCGGCGA 1922
 Db 1920 GGCACCTGTGGAAAGGCGCTCCAAAGCCCAACAAAGGAGTCAATCAAGATCGGGGA 1979
 Qy 1923 TGTCTATGTGAGAAACATTCAGAGGCAATGAAGCACTGGCGGCTCACCCTGTGAAGCAAG 1982
 Db 1980 CATCTGTCTCAGGAAATGGCCAGTTAAAGAGTTTACAGCTACTTCCAAAGACATGA 2039
 Qy 1983 CGAGGCTTGAAGCCCTGGAAGATGGAATCAAGAGCTCCCGGCGGTGGAAGACTTCTG 2042
 Db 2040 CGAGGTCTTAACGAATGGAAAAGGCTACCAAGCTGTGAAGATTTGAAGGCAAGTGA 2099
 Qy 2043 CAGAGACTTTAGACTCAGAAAGGTGTATACCTACCGCTCAACACCTTCTCTGCGGCC 2102
 Db 2100 CAAGAGTTTGAAGCTCAGAAAGGTGTGTACTTGCTCTCAACAGCTTCTGTGAAAGCC 2159
 Qy 2103 ACTGCAACGGGCTATGCACTACAGAGAGGTCTTGAGCGGCTGTGMAACACACCCGCC 2162
 Db 2160 CATCCAGCGGCTGTCTCACAACCGCTGTCTGTCGCCCGCTATGTGGACATTTACAGCCC 2219
 Qy 2163 GAGCCACGCGCACTTACAGGACTGCCGAGCCGCTTGTGGAGAGATCAGAGATGTGTGCG 2222
 Db 2220 CGGGCACATGACTAGCTGACTGCTGATACGCGCTTGAAGGCAATCAGAGGTGACAC 2279
 Qy 2223 ACAGCTCCACGATGATGATCAAGATGAGAAATTTCCAGAGCTGACAGAACTCAAGAA 2282
 Db 2280 CACACTACAGACATTTCTCATCCGGCTGGAGAACTGACAGAGCTAACGAGGTGACAGC 2339
 Qy 2283 AGATTTGATTTGGATTTGACAAATTTGTGTGTCGGGAAGAGGATTCATCCGTGCGGCG 2342
 Db 2340 GAGCTGTGTGGATGAGAACTCATGTGCTGTGCGAGGAGTTATCCGTGAGGGCTG 2399
 Qy 2343 CCTCAGCAAGCTTCCGGGGAAGGGGCTCCAGCAGCCCATGTTCTTCTGTTTCAACAGCCT 2402
 Db 2400 CTTTCAAGAGCTCACAAGAAAGGCTTGCAGCAGAGAGATGTTTTTCTGTCTTCAGATAT 2459
 Qy 2403 CCTGTATACAGAGCCGGGGGCTGACGGCTTCAATCAAGTTTAAAGTTCAACGGGAGCT 2462
 Db 2460 GTTGCTGTACACMAAGAGAGTTGACAGGAGCAAGCCACTTCCGATCCGGGGCTTCT 2519
 Qy 2463 CCGGCTTATGAGATGCAATTTGAGAGAGCAAGAGAGTGGGGGTGCCCACTGGCT 2522
 Db 2520 TCCCTTCAAGAGATCTGTGTGAAAGATGTAACGAGTGTCTGTTCACACTGTTT 2579
 Qy 2523 GACCCCTCCGGGGCAAGCGGAGTCCATCTGTCGCGGCAAGTTCTCGGTCCGAGATGA 2582
 Db 2580 CACCATCTACGGGCTCAGAAAACATCTGTGTGAGCAGCAGCACTCCGCTGAGAAA 2639
 Qy 2583 GAATGTGGTTGAGAGATCCAGATGGCCATTTGACTGTGGGAGAGAGACAGACCCCGC 2642
 Db 2640 GAATGTGATGTGAGCTGAATCTCCGATCCAGAGCCAGAGATGGGGGTGACACGGC 2699
 Qy 2643 CCTGTGTTCTGTGGCAGCA-----GCCCTGTGACAAAGTCCCTGATGAAAGCCAC 2696
 Db 2700 CCTGTGACTGCAAGCGGCGGCACTGTGTGCACTGTCCCCCAAGTATCCCAACAGGATATC 2759
 Qy 2697 CGCGGCTGACAGAGATCAGAGATGACCTGAGCGGCTCGCGCACTCGCTGAGAGGCA 2756
 Db 2760 TCTG---GAGCAGAGATTCAGAAATGATGTCTGGGGGTGTCGAGAGTCTCTGAGGGGGA 2816
 Qy 2757 GGGCCGCAACCGGCGCAACAAATGTGCACTGTGCTGTGGACCGCAACACAGCTTCT 2816

Db 2817 TGGCCAGACCGGGGCCAAACCAACAAATGCACTGTGTGTGATCCGGAACACCGAGCTGTC 2876
 Qy 2817 CATGTGACCTTACCATTCGACATGAGTAATTCATGTGCTGGAAACCTGTGAGAAAT 2876
 Db 2877 CAGGCGAGACCAATGTGACAGCTGTGAGAAACCAAGCTTCAAGATATCTGTAAAGAAATT 2936
 Qy 2877 CAAAACAGCAACCGGTGGCAGAAAGCTGTGGGTGTGTTTCAAACTTCTGCTTCTT 2936
 Db 2937 CAAAACAGTCAATGCTGGCAGAAAGCTGTGGGTGTGTTTCAAACTTCTGCTTCTT 2996
 Qy 2937 CTACAAATCAACACAGAACATATCCCTTGTGCAAGCTGCTGTGCTGCTACTCGCT 2996
 Db 2997 CTACAAATCAATCAATGATGATCTACCACTGGCCAGGCTCCGCTGTGCTGCTGAGAGCT 3056
 Qy 2997 CACCATCCCTCTGTGCTCGAAGATCCAGAAACCTAAGCTTCAAGCTGCTACTGCA 3056
 Db 3057 GAGCATCCCGAGGAGGCGGATGAGATTCACAAAGATGATGTTTCAAGCTCAGATTCA 3116
 Qy 3057 GTCCACGCTCTACTACTTCAAGGCGGAAAGCAGATACAGTTGAAAGGTGATGAAAGT 3116
 Db 3117 ATCCACGCTTACTTCTTCCGGGCTGAGAGCAAGTACATTTGAAAGGTGATGAGGT 3176
 Qy 3117 GATCCGACAGTGCACACAGCTGTGCTGCGGACCCC 3151
 Db 3177 GATCCAGGGGCGCAGACAGCTCAGCCGGAGGGCCC 3211

RESULT 13
 BC009153
 LOCUS
 DEFINITION
 Mus musculus PERM, Rhogef and pleckstrin domain protein 2, mRNA
 (cDNA clone MGC:6304 IMAGE:2655209), complete cds.
 BC009153
 VERSION
 KEYWORDS
 MGC.
 Mus musculus (house mouse)
 Mus musculus
 SOURCE
 ORGANISM
 Mammalia; Eutheria; Eumarchontia; Glires; Rodentia; Sciurognathi; Muridae; Muridae; Murinae; Mus.
 1 (bases 1 to 3719)

REFERENCE
 AUTHORS
 Strusberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Schenman,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schenfer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Ueda,T.B., Toshikiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullaly,S.J., Bosak,S.A., McMan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Halys,S.W., Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahy,J., Helton,E., Kettelman,M., Madan,A., Rodriguez,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,B.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalka,U., Smillie,D.E., Scherch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
 Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

JOURNAL
 PUBMED
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgabbe-remail.nih.gov

Tissue Procurement: Gilbert Smith, Ph.D.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Baylor College of Medicine Human Genome
 Sequencing Center
 Center code: BCM-HGSC
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
 Contact: amg@bcm.tmc.edu
 Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulsegod, H.,
 Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Navevati,
 A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
 Series: IRAC Plate: 5 Row: m Column: 20
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA g1: 21704021.

Location/Qualifiers

1. 3719

source

/organism="Mus musculus"
 /mol_type="mRNA"
 /strain="FVB/N-3"
 /db_xref="taxon:10090"
 /clone="MGC:5304 IMAGE:2655209"
 /issue_type="Mammary tumor, MMTV-LTR/INT3 model. 5 month
 old mouse. Taken by biopsy."
 /clone_lib="NCI CGAP_Mam2"
 /lab_host="DH10B"
 /note="Vector: PCMV-SPORT6"
 1. 3719

gene

/gene="Farp2"
 /note="synonym: Flr"
 /db_xref="GeneID:227377"
 /db_xref="MGI:2385126"
 99. 3296

CDS

/product="FERM, Rhogef and plectretin domain protein 2"
 /protein_id="AAH09153.1"
 /db_xref="GI:14318719"
 /db_xref="GeneID:227377"
 /db_xref="MGI:2385126"
 /translation="MGSEIGTYRALPTSGRLGGOTAGVTLPEPQSLSPMOEKM
 RIRVYKLDSTVELFDIEPKDGOVLNWKHNLTECYFGLEPQNSYWKPEPM
 KPIROYRRKRNALVRLAYKFPDPDQLOEYTRYLPAOLKRLLEBRLTTANTA
 ALIISLQSEIGDYDETLDRHLKANEYLPNOEKSLEKILDRHRTGQTPAESDFQ
 VLEIAKLIEVYGRFPMASDREGTKINLAVSHGVLPQGTINTFNMSKPKSPK
 RKRFLLKHPREVGPYODTLEFLGSRDECKFMKICVEYHFFRLSDPKPKAKAVF
 FSGSSFRSGTKQKOLVDYKDGKMKRIPEYRHSKTRTSLHALTVDPKOSVTD
 GLRTASLSANVSFYPRESSLSPGLPVLKDSLSLVDPQAPVTKSTAAEBSGAS
 SSGSPSTOSAHLPBPVLRPGPFSMDSPOPSLSKSHSLDCEPLQALSTRQAS
 PVLSPVLSAGTARMNOEQRKHMPEDEAVFAIEILATEKTYLKLDEVTWERS
 VILKEAMPALMALFNSNDIPVEYRFGHLEQRLALMEBPSAHKGDHQRIGD
 ILLRNROKLEFYSFQPHDEVLTELEKATKCKLEAVKEBELQKVCILPLNTPL
 KLPORLVHVRILSLCAHSPGRADYDEALKAITEVTELOOSLSTLENTLOLT
 ELORDLVGVEVILABREPTREGCLKTKGLOQMFPILFPMILYTSKTYGAHF
 RIRGFLPLKMLVESEBNEKSVLACFTTAAQKTIYVAASTRIEKKHMODLMAIOA
 AKTIQSPPLVGGPVYTRTPRSSDEVSLSESDGKNGSLGNSQHNATHYVCA
 YNTSVSRADHSAVENQLSGYLARKFKNSGKMGKLVVFTNCELFYANTHODYLA
 SLELGVSVSLPREADSIHNDYVKLOFQSHVFFPAESKYTERFWMVDVYKRASSSPG
 RPSFTQDQSHHSPLGEAIREKEACPSPCLDKNL"

ORIGIN

Query Match 29.6%; Score 1019.2; DB 6; Length 3719;
 Best Local Similarity 59.1%; Pred. No. 3.2e-278;
 Matches 1867; Conservative 0; Mismatches 1263; Indels 30; Gaps 6;

QY 70 CCGAGCCCGAGATCAGACTGGGGGCCCCGGAAATTCGGGATCGTACTTGGAACTG 129
 DB 132 CCAACTTCAGGACACGCTGGGTGGCCAGACCGCATTTGGAGTGAATCCCTGGAGCCA 191
 QY 130 GGACAGAGCCGCCCAACACCTTCAGAAAACCTCGTGTCCATCAAAATCCAGATGCTG 189

DB 192 GAGCAGAGCCTGTGACCCAGATGACGAGAAAGACATGCGCATCAGAGTGAAGCTGCTG 251
 QY 190 GATGACACCCAGAGGATTTGAAAGTCCAAAGAGCTCTGGGAAGTGTGCTGTGAT 249
 DB 252 GACAGACCTGTAAGTATTTTGCATCGAGCCGAACTGACGGGAGGCTTACTGACT 311
 QY 250 GCAGTTTGCAACCACTCAACCTCGTGAAGGATGATATTTGGCTCGAGTTTCTGTAT 309
 DB 312 CAGGTGTGAAGATTTGATATCTCATTTGAATGATGATCTATTTGTGTGAAGTTCAAGAAAT 371
 QY 310 CACAAAAGATCAAGGTGTGCTGATCTCTPAAACCATTTGTGAACAGATTTGAAGG 369
 DB 372 GTCCAGTCTTACGATTTGGCTTGAACCTATGAACCATCTATGAGCAATGACAAAG 431
 QY 370 CCAAGACAGTTTGTGTTAAGTTGTGTGAATTTCTTCCGCTGACACACACATC 429
 DB 432 CCAAAAAGCGGTGCTTGCCCTGGAGTAAATTTTCCCGCTTACCTGTGTCAGCTG 491
 QY 430 CAAAGAACTCAAGTATCTGTTGCGCTGACAGTGAAGCAGACTTGGCTCAAGGC 489
 DB 492 CAAAGAGTACAAAGTACTGTTTGTCTTGCATCAAGAGAGACCTCTGGAGAA 551
 QY 490 AGGTGACGTGTATGACACAGCCGAGCTCTTGTATTTACATATGTGCAATGTGAG 549
 DB 552 CGCTGACCTGACCGGCCAACCTGACGCCCTTCTCATATCCACTTGTGACGTGGAA 611
 QY 550 ATTGGGGAATTTGATGAAGCTTTGACAGAGACATTTAGCAAAAATAATATCATCT 609
 DB 612 ATCGAGATTAATGATTAACCCCTGATTCAGAAACACTCAAAAGCCATGATATCTTGCCC 671
 QY 610 CAGCAAGACGCACTAGAGCAAAATCGTGAATTTCAACATTAACCAATTTGAGCAAA 669
 DB 672 AACCAAGAAATCTTGAAGAAAGATCTAGACTTCAATCAAGAGCACAGGCCAGACT 731
 QY 670 CCAAGCAATCAAGATTTCCAGCTCTAGATATGCGCGTGGCTAGAGATGATGGAATC 729
 DB 732 CTGACAGATCAGATTTCCAGGCTTGAATGACAAAGAGCTGGAATATGATGCAATC 791
 QY 730 CGGTTCACCCGCGCAAGACAGGAAAGGACGAAGATCAATTTGGCCGTTGCCAACAG 789
 DB 792 AGGTTTCAATGCTTCTAACAAGAAAGGACCAAAATTAATCTAACAATTTCTCAATG 851
 QY 790 GGAATCTAGTGTTCAGGATTTCACTAAGATCAATGACCTTCAACTGGGCAAGGTGCG 849
 DB 852 GGTGTCTGTGTTCAGGATACACCAAAATCAATCTTCACTGCTGCCAAGTCCGT 911
 QY 850 AAGCTAGCTTCAAGAGAGCGCTTCTCATCAAGCTCGGCAAGTCCATATGTCG 909
 DB 912 AACCTAAGCTTCAAGGAAAGGTTTCTTATCAAACTCAACCTGAGGTCCATGGGCC 971
 QY 910 TACCAAGATCTTGAATTTCTGATGCGCAGTGGGATTTCTGCAAGTCTTTTGAAA 969
 DB 972 TACCAAGACAGTTAGATTTCTGTGGGTAGAGAGAAATGTAAGAACTTGTGAAAG 1031
 QY 970 ATCTGTTGAACATGATGCTCTTCTTGAAGAGGAGCCAAAGCCAAAGCCAAAG 1029
 DB 1032 ATATGTGTGAAGTACATCTTTTAACTCTGACCAAGCTTAAAGCAAGCAAG 1091
 QY 1030 CCGTCTCTTTAGCCGGGAGTATCATTTCCGTTCAAGTGTGGAATCTGAGAGCAGTT 1089
 DB 1092 GGTGTCTTCTTACGCGGAGGCTCTCTCAATACAGTGAAGAACTCAGAAACAATCA 1151
 QY 1090 CTGACATATGTAAGAGAGAGACATTAAGAAAGTGCAGTTGAAAAGAGACACAGCAAG 1149
 DB 1152 GATGATTAATGTAAGACGATGTAAGAAAGATTTCCATAGAAAGAGGACACATTAAG 1211
 QY 1150 ATTC--ATTCTATCCGAGGCTTGTGCTTACAGCTTACAGACTGAATTCGGAAGTCTG 1206
 DB 1212 ACTCGACATCTCTTATATCTGTAGTATGATCTGCTTAAACAGAGCTCTCTTACAC 1271
 QY 1207 GAGCAGTCTCAAGACAGACAGCCTTACATTTGAGAAAGGTGCGCAATCTCCAGGGGCG 1266
 DB 1272 GATGGCTGAGAACTTCTGCTCTGCTTCAAGCAAAATGTCTCTTATATCC----- 1324

QY 1267 CAGAGTCGCGGAGAGAAAGAACCGAAGTTTCCGCGGAGGCGGAGTCCGACCCG 1326
Db 1325 ---ACCCTTAAGTTCTTTTGTCCCTCTGCTGCTGCTTGAAGAACAGACGACG 1381
QY 1327 AGCCCTGCGCGAGAGAAAGCCCGCGGGTAAACAGACGCGAGCGAGCCGCTCGCG 1386
Db 1382 CTCCTTGTGATTCGCGAGGCTCCCGTCAATCAAGACAGACGAGAGAGGAGCGCG 1441
QY 1387 CCCACGAGAGAGAGAGAGGTCTTAAGATAGACCCAGCAGATTAACCTCAGCCC 1446
Db 1442 ACCA--TCATCATCTTACGCGCCCGACACAGTCGCGCCATCTCCCTGACCCCGGTG 1499
QY 1447 CCGCAGCCAAAGACAGAGTCCCTGACTGGGAGCTCCATCCTTTCCGAGCTGTCTGTAAC 1506
Db 1500 CTCGCGCTGTGTCAGAGCTTTTCTATGATGATGCTCTCAGCTTCTCCCTCAGCCGTAAG 1559
QY 1507 TCGCAGAGGGAGAGTGGCCCTGCGCAAGTGAAGTGTCTCCAACTGAGCCCGACACC 1566
Db 1560 AGCAGCTGAGGCTGTGCTGAGCTTCAAGCCGCG-----CTGAGTACAGCTGAG 1610
QY 1567 AAGCAGGCTCTTCCCTTGAATCAGCCCGCTGTGAATGACAGGCTCGCCCGGAGAGAC 1626
Db 1611 CAGGGTGAATCCCTGTGTCTAGCCCTGTGCTCAGGCGCTGTGACAGCCAGATGAGAC 1670
QY 1627 GATGAGATGAGGGCCGAGAGAAAGATTCCTCAATGATAAAGCGTACTTCAATAGCTAAG 1686
Db 1671 AACCAAGAGAGAGAAACACAGACATGCGCAAGAGATGAGGCTTATTCATAGCAAG 1730
QY 1687 GAAGTGTCTACCAACGAGCGAAATATCTGAAGATCTCGAAGTTATCACTTGTGTTT 1746
Db 1731 GAGATCTCGCTACAGAACGAACTTATGAGAGATTTAGAGTTATCACTGTGTGTTT 1790
QY 1747 CAGAGCAGGTGAGCAAAAGAGACCGCATGCGGAAAGCACTGAAAAGTCTCATTTCCG 1806
Db 1791 CGAGAGTGTCTATCAAGAGAGAGGAGGCTGCGAGCCCTGATGAGCCCTGCTTTCTCC 1850
QY 1807 AATTTGAACCTTTGACAAATTTCACTAATTTTCTCAAGAAATAGACACGACTT 1866
Db 1851 AACATTGATTCAGTCTACAGATTCACAGAGCTTCTTCAAGAGTGAACAGAGGCTG 1910
QY 1867 GCCCTGTGGAGAGGCGCTCAATGCGCA--AATCAGAGATTAACCAAGAAATCGCGAT 1923
Db 1911 GCACTCTGGAGAGGCGCTCAAGTGCCTTAAAGGTGATCAACAGGAAATCGGGAGAC 1970
QY 1924 GTCATGCTGAAGAACTTCAAGGCAAGAACCTGCGGCTCACTGTGGAAGACACAGC 1983
Db 1971 ATCTCTCTCAGAGAACTGCGTCACTTAAAGAAATTTAAGCTTATCCAAAGACAGAT 2030
QY 1984 GAGGCTTTGAGGCGCTTGAAGAAATGGAATCAAGAGCTCCCGGCGCTGAGAACTTCTGC 2043
Db 2031 GAGGTCTTAACAGAACTGGAAGAGGCAACAAACATGTAAGAAAGTGAAGGCAAGCTAC 2090
QY 2044 AAGAGCTTTGAGCTGAGAAAGGTGTGTAACCTACGCTCAACACTTCTCTCGCGCCA 2103
Db 2091 AAAGAGTTTGAAGTCCAAAGGCTCTACCTGCTCAACACTTCTCTGTAAGGCC 2150
QY 2104 CTGCAACGGCTCATGACTACAAAGAGTCTGAGACGGCTGTGCAACACACCGCGCG 2163
Db 2151 GTCCAGAGGCTGAGTCACTACGCTGTGCTGAGCGGCTGTGCTCACTACTCTCT 2210
QY 2164 AGCCACGCGGACTTCAAGGACTGCGAGCGCTTTGGCAGAGATCAACGAGATGTGTGCA 2223
Db 2211 GGGCACCAGGACTATGCCGACTGCGACTGAGGCACTGAAGCCATCAACAGAGACCAAC 2270
QY 2224 CAGCTTCACGAGTACGATGATCAAGATGAGAAATTTCCAGAGCTGACAGAACTCAAGAAA 2283
Db 2271 GAGCTCACACAAAGCTTACCGGCTGAGAAACCTTCAAGAAATTTGACGAGCTACAGCA 2330
QY 2284 GATTTGATTTGGCATTTGACATCTGTGTGTTCCGGGAGGAGATTCATCCGTGCGGACAGC 2343
Db 2331 GACCTGTGAGGTGTAAGAAACCTCATTTGCTCTGGAGGAGATTTATCCGTGAGGCTGC 2390

QY 2344 CTCAGCAAGCTCTCGGAGAGAGGCGCTCCAGACAGCATGTTCTTCTGTTCAACGACGTC 2403
Db 2391 CTGCAACAAAGCTACCAAGAGGCGCTGCAACAGAGAGATTTTTTCTGTTCTCAGATATG 2450
QY 2404 CTGCTATACAGAGCCGGGGGCTGACGCTTCATCACTTTAAAGTCCACGCGGACGTC 2463
Db 2451 TTACTGTATACAGCAAAAGTGTCAACAGAGCCAGTCACTTCCGAGTCCGTGGCTTCTT 2510
QY 2464 CCCCTCTATGAGATGAGATTGAGGAGCGAAGACGATGGGGGGTCCCCACTGCGCTG 2523
Db 2511 CCATCCGTGGATGCTGTGTAAGAAAGTGAATGATGATGATGATGATGATGATGATGATG 2570
QY 2524 ACCCTCCGGGGCAGCGGACGATCCATCATCTGTGCGCGCAGTCTCTGATCCGAGATGAG 2583
Db 2571 ACCATCTATGACGCTCAAAAAACATTTGTGTAGACGACGACACTCGGCTAGAAAAAGAA 2630
QY 2584 AAGTGGTTGAGAGATCAATGAGCCATTGAACCTGGCGGAGAAAGACAGACGCCCGCC 2643
Db 2631 AAGTGAATGACAGACCTTAATGACAGCAATCCAGACGCAAGACTATGCTGATCTACCC 2690
QY 2644 CCTGAGTCTTGGCCAGGACCCCTGACAAAGTCCCTGATGAAGCCACCGGCT 2703
Db 2691 CCAAGTCTGCTGG--GAGGCCGGGTATACCTGATACCCCTAGATCTTGTGATGAATG 2747
QY 2704 GACCAAGATGACAGATGACCTGAGCGCTCGCGCAATCGCTGAGACGCCAGCCCG 2763
Db 2748 TCTGTGAAGAAATCAAGAAATGATGTGAGAAACCGGGGCTCCCTGAGGGGAAACAGCCAG 2807
QY 2764 CACCGCGGCAACAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2823
Db 2808 CACCGGGCAATACAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2867
QY 2824 GACTTACAGATCGCAGTGAAGATCAGTGTCTGGAACCTGCTGAGGAAATTCAAAAAC 2883
Db 2868 GACCAAGTGAAGCTGTGTAAGAACAGCTTTCAAGATATCTGTGTAAGAAATTCAGAAAC 2927
QY 2884 AGCAACGGGTGACAGAGCTGTGAGTGTGTTCAAACTTCTGCTGTTCTTACAAA 2943
Db 2928 AGTATGCTGGCAGAACTCTGGGTGCTTTTACCACTTCTGCTGTTCTTATATAA 2987
QY 2944 TCACACAGACATATCCCTTGGCCAGCTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3003
Db 2988 ACACATCAAGATGATTAACCCCTGCGCAGCTCCCACTATGGTTACAGCGTGAAGCTC 3047
QY 3004 CCCCTGAGTCCGAGAACATCCAGAAAGACTACGTTCAAGCTSCACTTCAAGTCCAC 3063
Db 3048 CCCAGGAGGCTGACAGATTCACAAAGACTATGTTCAAGCTTCAATTCATCTCAC 3107
QY 3124 AGTCCACCAAGTCTGTGCTCGGACCCCAAGTGTGAGCCACAAAGATCTTGTGAT 3183
Db 3168 AAGGCAACAGCTCAACAGGAGACCCCAAGTTCACTCAAGAGCTGCTCAATCATCTCT 3227
QY 3184 TGATGCGCGACACACTGTTCCGAGTGTGCTTCC 3223
Db 3228 CCAAGGCTGAGGACAGATCAAGAAAGAGGATGCC 3267

RESULT 14
BC027077 2431 bp mRNA linear ROD 21-OCT-2003
LOCUS BC027077
DEFINITION Mus musculus PERM, RhogEP (Aryge) and pleckstrin domain protein 1
(chondrocyte-derived), mRNA (CDNA clone IMAGE:5010682), partial
cde.

ACCESSION BC027077 GI:20071584
VERSION BC027077.1
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE
AUTHORS
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathia; Muridae; Muridae; Murinae; Mus.
1 (bases 1 to 2431)

Strauberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heide, F.,
Datchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Scapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Scheetz, J.E., Brownstein, M.J., Uedl, T.B., Toshlyuk, S.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalins, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932

TITLE
JOURNAL
PUBMED
REFERENCE
AUTHORS
JOURNAL
TITLE
Direct Submission
Submitted (04-APR-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

REMARK
COMMENT
NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk
Email: cgabs-r@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: amg@bcm.tmc.edu
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulsaged, H.,
Kowit, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNLN at: <http://image.llnl.gov>
Series: IRAC Plate: 45 Row: C Column: 12
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency OR
analysis.

FEATURES
source
Location/Qualifiers
1. 2431

/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="FVB/N-3"
/db_xref="taxon:10090"
/clone="IMAGE:5010682"
/clone_type="Mammalian tumor, MMTV-LTR/INT3 model, 5 month
old mouse, Taken by biopsy."
/clone_1lb="NCI CGAP_Mam2"
/lab_host="DH10B"
/note="Vector: PCMV-SPORT6"
1. 2431
/gene="Farpl1"
/note="synonym: Cdap"
/db_xref="GeneID:223254"
/db_xref="MGI:2446173"
1. 1065
/gene="Farpl1"
/cdon_start=1

/product="Farpl protein"
/protein_id="AAH27077.1"
/db_xref="GI:20071585"
/db_xref="GeneID:223254"
/db_xref="MGI:2446173"
/translation="YKGVLEKCKKHPPHNADFRCDRAALAEITENVAQLHGTMTME
NQKHEHLKDLIGDNLVTPGSEFRLQSLKLSKQIQOMFFENVLVLTSGE
TYSNPKVGOLPLVGMTIEESEEEVGHCLTRQROSIIVASRSEMEKMEKI
QMAIDAEKNGPTELLASPPDNDSPDRAODESDIISASTISERAPRPHG
TWHVWHRSTSVSWDESIANVENQSGNULKFRKSNQKMLVFTNFCPEYKSH
QSHPLASLPILGYSLTIPSESEENHKDVFKLHFKSHVYFRAESEYFPERMEVIR
SATSSASRAHLIISHKESHLV"
250. 501

/gene="Farpl1"
/note="PH: Region: Pleckstrin homology domain. Domain
commonly found in eukaryotic signalling proteins. The
domain family possesses multiple functions including the
abilities to bind inositol phosphates, and various
proteins. PH domains have been found to possess inserted
domains (such as in PLC gamma, syntrophins) and to be
inserted within other domains. Mutations in Brutons
tyrosine kinase (Btk) within its PH domain cause X-linked
agammaglobulinemia (XLA) in patients. Point mutations
cluster into the positively charged end of the molecule
around the predicted binding site for phosphatidylinositol
lipids"
/db_xref="CDD:smarr00233"
730. 1020

/gene="Farpl1"
/note="PH: Region: Pleckstrin homology domain. Domain
commonly found in eukaryotic signalling proteins. The
domain family possesses multiple functions including the
abilities to bind inositol phosphates, and various
proteins. PH domains have been found to possess inserted
domains (such as in PLC gamma, syntrophins) and to be
inserted within other domains. Mutations in Brutons
tyrosine kinase (Btk) within its PH domain cause X-linked
agammaglobulinemia (XLA) in patients. Point mutations
cluster into the positively charged end of the molecule
around the predicted binding site for phosphatidylinositol
lipids"
/db_xref="CDD:smarr00233"

ORIGIN

Query Match 25.8%; Score 889.6; DB 6; Length 2431;
Best Local Similarity 89.0%; Pred. No. 2,6e-241;
Matches 961; Conservative 0; Mismatches 119; Indels 0; Gaps 0;
QY 2122 TACACAGAGCTCTTGAGGAGGCTGTGCAACACACCCCGAGCCAGCCGACTTCAGG 2181
DB 1 TATACAGATGCTCTTGAGAGGCTGTGCAACACACCCAGCCAGCCGACTTCAGG 60
QY 2182 GACTGCCAGGCGCTTTGGCAGAGATCACGAGATGTTGGCAGAGCTTCAGGTTACGATG 2241
DB 61 GACTGCAGAGCTGCGCTGGCGAGATCACAGAGATGTTGGCAGAGCTTCAGGTTACGATG 120
QY 2242 ATCAAGATGAGAAATTTCCAGAAAGCTGCAGAACTCAAGAAAGATTTGATTTGCAATTGAC 2301
DB 121 ATCAAGATGAGAAATTTCCAGAAAGCTGCAGAACTCAAGAAAGATTTGATTTGCAATTGAC 180
QY 2302 AATCTTGATGTCGAGGAGGAGGATTCATCCGCTGGGAGAGCTTCAGAAAGCTTCGAGG 2361
DB 181 AATCTTGATGTCGAGGAGGAGGATTCATCCGCTGGGAGAGCTTCAGAAAGCTTCGAGG 240
QY 2362 AAGGGGCTCAGAGGAGGATTTCTTCTGTTCAAGCAGAGCTCTGTTATACAGAGCCGG 2421
DB 241 AAGGGGCTCAGAGGAGGATTTCTTCTGTTCAAGAGATGTTCTGTTATACAGAGCCGG 300
QY 2422 GGGCTACGCGCTTCACATAGTTTAAAGTCCAGCGGAGAGCTCCGCTTATAGGCAATGAG 2481
DB 301 GGACTACGCGCATATATAGTTTAAAGTCCAGCGGAGAGCTCCGCTTATAGGCAATGAG 360
QY 2482 ATTGAGAGAGCGAAGACAGATGAGGAGGAGTCCCACTGCTGACCTTCGAGGAGCAGCG 2541

Db	Accession	Source	Definition	Length
Db	361	ATCGAGGAGATGAGAGAGAGATGGGGGTGTGCCCCCACTGCTGACCTTGTGGGGCCAGCGG	420	
Qy	2542	CAGTTCATCATCTGTGGCCGCCAGTTCTCGGTCCGATGAGAAAGTGGGTTGAGACATC	2601	
Db	421	CAATTCATCATCTGTGGCTGTCCAGTTCCCGGTGAGAGATGGAATAATGATGAGAGACATC	480	
Qy	2602	CAGATGGCCATTGACCTTGGCCGAGAAAGCAGACAGCCCCGCCCTTGATTTCTTGGCCAGC	2661	
Db	481	CAGATGGCCATTGACCTTGGCCAGAAAGCAATGAGCCACCCCCGAGCTCTTGCCAGC	540	
Qy	2662	AGCCCCCTTGACCAACGAATGCCCTGATGAAGCAGCGGGCTGACAGAGATCAGAGGAT	2721	
Db	541	AGCCCCCTTGACCAACGAATGCCCTGATGAAGCAGCGGGCTGACAGAGATCAGAGGAT	600	
Qy	2722	GACCTGAGAGGCTCTGCGCAGACATCGCTGAGAGCGCCAGGCCGCCACCGGGCAACAATG	2781	
Db	601	GACCTGAGAGGCTCTCTCGGACATCGCTGAGAGCGCCAGGCCGCCACCGGGCAACAATG	660	
Qy	2782	GTGCACGTGTCTGTGACCGACCGCAACCAAGCTTCCATGTGAGATTGACATCGCAGTG	2841	
Db	661	GTGCACGTGTCTGTGACCGACCGACCAAGCTTCCATGTGAGATTGACATTTGCTGTG	720	
Qy	2842	GAGAAATCAGTTGTCTGAGAACTCTGTGAGAAATTCAAAACAGCAAGGGTGGAGAG	2901	
Db	721	GAGAAACAGCTCTCTGGGAACTTGTGAGAAAGTTCAAGAAACAGCAAGGGTGGAGAG	780	
Qy	2902	CTGTGGGGTGGTGTTCACAAACTTTCGCGCTTCTTCTPAACAATCAGACAGGACAAATAT	2961	
Db	781	CTGTGGGGTGGTGTTCACAACTTTCGCGCTTCTTCTTCAAAATCAGACAGGACAGTAC	840	
Qy	2962	CCCCCTTGGCCAGCTGCTCTGTCTGCGCTACTCGCTCACCATCCCTCTGATGCCAGAAC	3021	
Db	841	CCCCCTTGGCCAGCTGCTCTGTGTGGGCTTACTGCTGCTTCTTCAAAATCAGACAGGAC	900	
Qy	3022	ATCCAGAAAGACTACGTGTTCAGCTGCACTTCAAGTCCACGTTACTTACTTCAAGGCG	3081	
Db	901	ATCCAGAAAGACTATGTGTTCAGCTGCACTTCAAGTCCACGTTACTTACTTCAAGGCG	960	
Qy	3082	GAAAGCGAGTACAGCTTGAAGAGGTGAGAAAGATCCGACAGGCCACACAGCTCTGCG	3141	
Db	961	GAAAGCGAGTACACTTTTGAAGAGGTGAGAAAGATCCGACAGGCCACACAGCTCTGCG	1020	
Qy	3142	TCCGACACCCCACTGTGAGCCACAAAGATCTCTTGTATGATGAGCCGACACATC	3201	
Db	1021	TCTGGGGCCCACTTGTGAGTCAAAAGATCTCACTGTACTAGCGCTGACACGCGC	1080	
RESULT 15	BC043327	2296 bp	mRNA	linear
LOCUS	BC043327			
DEFINITION	BC043327			
KEYWORDS	BC043327			
ACCESSION	BC043327			
VERSION	BC043327.1			
KEYWORDS	GI:28175181			
SOURCE				
ORGANISM				
REFERENCE				
AUTHORS				

REFERENCE
AUTHORS

1 (bases 1 to 2266)
Strausberg, R. L., Feingold, E. A., Grouse, L. H., Derge, J. G.,
Klausner, R. D., Collins, F. S., Wagner, L., Shenmen, C. M., Schuler, G. D.,
Altschul, S. F., Zeeberg, B., Bustow, K. H., Schaeffer, C. F., Bat, N. K.,
Hopkins, R. F., Jordan, H., Moore, T., Max, S. I., Wang, J. F., Hsieh, F.,
Datsenko, M., Marwala, K., Farmer, A. A., Rubin, G. M., Hong, L.,
Stapleton, M., Soares, M. B., Donald, M. F., Casavant, T. L.,
Scieczek, T. E., Brownstein, R. J., Ueda, T. B., Toshitsuki, S.,
Canciari, P., Prange, C., Rama, S. S., Loquiano, N. A., Peters, G. J.,
Abramson, R. D., Mullany, S. J., Bosak, S. A., McEwan, P. J.,
McBarnett, K. J., Malek, J. A., Gunaratne, P. H., Richards, S.,

TITLE	JOURNAL
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Wallison, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Rahcy, J., Helton, E., Ketteman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakeley, R.W., Touchan, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmitt, J., Myers, R.M., Buckfield, Y.S., Krzywinski, M.I., Skalske, U., Smalutz, D.E., Scherer, A., Schein, J.E., Jones, S.J., and Marra, M.A. Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences	PROC. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
JOURNAL PUBLISHED	12477932
REFERENCE	2 (bases 1 to 2296)
AUTHORS	Strausberg, R.
TITLE	Direct Submission
JOURNAL	Submitted (09-JAN-2003) National Institutes of Health, Mammalian

REMARK	COMMENT
NIH-MGC Project URL: http://mgc.ncl.nih.gov	
Contact: MGC help desk	

Tissue Procurement: Gilbert Smith, Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@bcsc.bc.ca

Steven Jones, Jennifer Asano, Ian Boeset, Yaron Butcherfield,
Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
Leticia Heiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
Sen Lee, Victor Ling, Carrie Matheson, Candice McLeavy, Steven
Ness, Pawan Pandoh, Anna-Lisa Pridhu, Parvaneh Speedi, Jacqueline
Schain, Duane Smalnu, Michael Smith, Lorraine Spence, Jeff Scott,
Michael Thorne, Miranada Tsai, Natascha van den Bosch, Jill Vardy,
George Yang, Scott Zuyderduyn, Matteo Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/BLN at: <http://image.llnl.gov>
Series: IRAK Plates: 86 Row: 0 Column: 23
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis.

FEATURES	Location/Qualifiers
source	1. .2296

gene

CDS

```

/product="Purp protein"
/protein_id="AAH43327.1"
/db_xref="GI:28175182"
/db_xref="GeneID:223254"
/db_xref="MGI:2446173"
/tranlation="KLHELPKDLIGDLVTPGREFIRGLSGSLKSGKQHQMPFLF
NDVLLATSGGLANQPKYQKGLPGYGMTIESSBEGVGHCLTGORSLIIVASS
RSEKRMEDIDQALDIAEKSGNPTELLASSPPDKSDDEATADQESDDLSAST
SLERAPFHGNTSVHCHRSTSVSMVDSIAVENQSLNLKFKNSNGMQLWVF
INFCLEFPHGSDHSHSLPLGLGSLTIPSESENTHKDYVFLGKSHVYFRASE

```

misc_feature
YPERMMEVIRSATSSASRAHILSHKESHLV"
111..362

/gene="Parp1"
/note="PH; Region: Pleckstrin homology domain. Domain commonly found in eukaryotic signalling proteins. The domain family possesses multiple functions including the abilities to bind inositol phosphates, and various proteins. PH domains have been found to possess inserted domains (such as in PLC gamma, syntrrophins) and to be inserted within other domains. Mutations in Brutons tyrosine kinase (Btk) within its PH domain cause X-linked agammaglobulinaemia (XLA) in patients. Point mutations cluster into the positively charged end of the molecule around the predicted binding site for phosphatidylinositol lipids"
/db_xref="CD:smart00233"
591..881

misc_feature

/gene="Parp1"
/note="PH; Region: Pleckstrin homology domain. Domain commonly found in eukaryotic signalling proteins. The domain family possesses multiple functions including the abilities to bind inositol phosphates, and various proteins. PH domains have been found to possess inserted domains (such as in PLC gamma, syntrrophins) and to be inserted within other domains. Mutations in Brutons tyrosine kinase (Btk) within its PH domain cause X-linked agammaglobulinaemia (XLA) in patients. Point mutations cluster into the positively charged end of the molecule around the predicted binding site for phosphatidylinositol lipids"
/db_xref="CD:smart00233"

ORIGIN

Query Match 22.6%; Score 777.8; DB 6; Length 2296;
Best Local Similarity 89.2%; Pred. No. 1.9e-209;
Matches 839; Conservative 0; Mismatches 102; Indels 0; Gaps 0;

QY	2261	AGAAAGCTGCACGAACCTCAAGAAATTTGATGGCATTCATCTTGTGTTCCGGGAA	2320
DB	1	AGAAAGCTGCATGAGCTCAAGAAATTTGATGGCATTCATCTTGTGTTCCGGGAA	60
QY	2321	GGGAGTTTCATCCGCTTGGGAGCCTCAGCAAGCTCTCGGGAAAGGGCTCCAGCAGGCA	2380
DB	61	GGGAGTTTCATCCGCTTGGGAGCCTCAGCAAGCTCTCGGGAAAGGGCTTCAGCAGGCA	120
QY	2381	TGTTCTTCTGTTCAACGACGTCTGTATACAGAGCCGGGGCTGACGGCTTCAATC	2440
DB	121	TGTTCTTCTGTTCAACGATGTCTGTGTATACAGCCGGGGATGACGGCATCTAATC	180
QY	2441	AGTTTAAAGTCCACGGGACGCTCCGCTATGGCANTGACGATTGAGGAAGGAAGACG	2500
DB	181	AGTTTAAAGTCCACGGACAGCTCCCACTATGGCATGACGATGAGAGATGAGGAGG	240
QY	2501	AGTGGGGGGTGGCCCACTGCTGACCTCCGGGGCCAGCGGAGTCCATCATGTGGCCG	2560
DB	241	AGTGGGGTGGCCCACTGCTGACCTTGGGGCCAGCGGCAATCATGTGGCTG	300
QY	2561	CCAGTTCTCGGTCGAGATGGAAGTGGTTGAGCATTCAGANTGGCAATTGACCTGG	2620
DB	301	CCAGTTCCCGGTGAGATGGAATAATGATGAGGACATTCAGATGGCAATTGACCTGG	360
QY	2621	CGGAGAAAGACAGACGCCCCCGCTGAGTTCCTGGCCAGACGCCCCCTGACAAACAAT	2680
DB	361	CAGAGAAACCAATGAGCCACCCCGAGCTCTTGGCAGACGCCCCCTGACAAACAAT	420
QY	2681	CCCTGTGTAAGCCACCGCGGTGACCAAGATCAGAGATGACCTTGAAGCGCTCGGCA	2740
DB	421	CCCTGTGTAAGCCACCGCGGTGACCAAGATGAGAGATGACCTTGAAGCGCTTCGGA	480
QY	2741	CATGCTGAGAGCGCCAGGCCCCGACCGCGGCAACACATGTTGTCAGTGTGCTGGACC	2800
DB	481	CATGCTGAGAGCGCCAGGCCCCCTCACCGTGGCAACACATGTTGTCAGTGTGCTGGACC	540

QY	2801	GCAACACACGCGTCTTCATGATGAGACTTCAGCATGCGACGTGAGATCAGTTGCTGAA	2860
DB	541	GCAACACACGCGTCTTCATGATGAGACTTCAGCATGCTGTGAGAAACAGCTCTGGGA	600
QY	2861	ACCTGCTGAGAAATTCAAAAACAGCACGGGTGCGAAGCTGTGGTGTGTTACAA	2920
DB	601	ACTTGTGAGAAATTCAAAAACAGCAATGGCTGGCAGAAAGTGTGGTGTATTCACCA	660
QY	2921	ACTTGTGCGTGTCTTCTCAAAATCACACGAGACAAATATCCCCCTTGCAGCGCTTC	2980
DB	661	ACTTGTGCGTGTCTTCTCAAAATCACACGAGACAGTACCCCTTGCAGCTGCTTC	720
QY	2981	TGCTGCGTACTGCTGACCATCCCTCTGAGTCCGAGAACATCCAGAAAGACTACGTG	3040
DB	721	TGCTGCGTACTGCTGACCATCCCTCTGAGTCCGAGAACATCCAGAAAGACTATGTG	780
QY	3041	TCAGCTGCACTTCAAGTCCACGCTTACTTCAAGGCGGAAAGCGATACGTTTG	3100
DB	781	TCAGCTGCACTTCAAGTCCACGCTTACTTCAAGGCGGCTGAAAGCGAATACACTTGG	840
QY	3101	AAAGGTGATGGAAGTATCCCGAGTCCACAGCAGCTTGGCGGACCCCACTGTTGA	3160
DB	841	AAAGGTGATGGAAGTATCCAGAGTCCACAGCAGCTTGGCGGACCCCACTTGA	900
QY	3161	GCCACAAAGAGTCTTGTGTATTTGATGGCCGGACACACTC	3201
DB	901	GTCACAAAGATCTCACCTGTACTGACGGCTGAGACGCGCC	941

Search completed: May 22, 2006, 07:58:51
Job time : 12710.4 secs

This Page Blank (user)

Aas27016 cDNA enco
Adb93194 Human cDN
Ach21582 Human adu

Adc71321 Human col

Ach47450 Human inf
Abx34991 Bovine ES
Aas64761 DNA encod

Aaa42150 Human sec

Adb93618 Human cDN

Ab190828 Human pol

Adp10335 Reference

Adp54688 Human PRO

Abv30214 Human pro

Adv43235 Human psy

ke domain; cancer;
atoid 'arthritis;

PF 27-NOV-1998; 98MO-JP005348.
XX
PR 27-NOV-1997; 97JP-00342060.
XX
PA (CHUS) CHUGAI SEIYAKU KK.
XX
PI Kato Y, Kawamoto T, Koyano Y;
XX
DR WPI: 1999-371117/31.
DR P-PSDB; AA07482.
XX
PT Protein CDEP expressed in differentiated chondrocytes, and gene encoding
XX it.
XX
PS Claim 5; Fig 1; 59pp; Japanese.
XX

This sequence represents the coding region for a protein (CDEP) expressed
CC in differentiated human foetal chondrocytes, which contains an ezrin-like
CC domain, a Dbl homology (DH) domain and a pleckstrin homology (PH) domain.
CC The nucleic acid or protein can be used in the investigation and
CC treatment of cancers and arthritic diseases (including chronic rheumatoid

CC arthritis), or for screening of candidate anticancer drugs
XX
SQ Sequence 3442 BP; 864 A; 952 C; 927 G; 699 T; 0 U; 0 Other.

Query Match	100.0%	Score 3442	DB 2	Length 3442
Best Local Similarity	100.0%	Pred. No. 0		
Matches 3442; Conservative	0	Mismatches	0	Gaps 0

OY	1	UGCCGCAAGCCCGCGCGCTGTGGAGATATTCTCTAAGCCGCTTTTCATCATGGAGAAATA	60
Db	1	CGCCGCAAGCCCGCGCGCTGTGGAGATATTCTCTAAGCCGCTTTTCATCATGGAGAAATA	60
OY	61	GAGCAGAGGCGGACCCCGAGATTCACGATGGGAGGCGCCGGAAAAATTCGGGGATCAGTACC	120
Db	61	GAGCAGAGGCGGACCCCGAGATTCACGATGGGAGGCGCCGGAAAAATTCGGGGATCAGTACC	120
OY	121	TTGGAACTGGACAGAAAGCCGCCCAACACCTTTCAGGAAAACTGTGTTCATCAAAATC	180
Db	121	TTGGAACTGGACAGAAAGCCGCCCCAACACCTTTCAGGAAAACTGTGTTCATCAAAATC	180
OY	181	CAGATGCTGGATGACACCCAGAGGCACTTAAAGTCCACAAABAAGCTCCCGGGAAAGTG	240
Db	181	CAGATGCTGGATGACACCCAGAGGCACTTAAAGTCCACAAABAAGCTCCCGGGAAAGTG	240
OY	241	CTGCTGATGAGTTTGCAACCACTCAACCTCGTGAAGGTGACTAATTTTGGCGCTGAG	300
Db	241	CTGCTGATGAGTTTGCAACCACTCAACCTCGTGAAGGTGACTAATTTTGGCGCTGAG	300
OY	301	TTTCTGTATCAAAAAGATCACGGTGTGCTGATCTCTTAAACCCATTGTGAACAG	360
Db	301	TTTCTGTATCAAAAAGATCACGGTGTGCTGATCTCTTAAACCCATTGTGAACAG	360
OY	361	ATTGAAGGCAAAAGCAGCTGTGTTGAAGTTTGAGTGAATCTTTCCGCGCTGACAC	420
Db	361	ATTGAAGGCAAAAGCAGCTGTGTTGAAGTTTGAGTGAATCTTTCCGCGCTGACAC	420
OY	421	ACACAACTCCAAAGAACTCAACAAGTACTGTTCGCGCTGAGGTGAACAGACAGCTTG	480
Db	421	ACACAACTCCAAAGAACTCAACAAGTACTGTTCGCGCTGAGGTGAACAGACAGCTTG	480
OY	481	GCTCAAGGCAAGTTGACGTGTATGACACCAAGCGCAGCTCTCTTGATTTACACATTGTG	540
Db	481	GCTCAAGGCAAGTTGACGTGTATGACACCAAGCGCAGCTCTCTTGATTTACACATTGTG	540
OY	541	CAATCTGATTTGGGGATTTTGAAGAACCTTGGACAGAGGCACTTAGCAAAAATAATA	600
Db	541	CAATCTGATTTGGGGATTTTGAAGAACCTTGGACAGAGGCACTTAGCAAAAATAATA	600
OY	601	TACATACCTCAGCAAGACGCACTAGAGGACAAAAATCGTGAATTTCAACCATMACCATT	660
Db	601	TACATACCTCAGCAAGACGCACTAGAGGACAAAAATCGTGAATTTCAACCATMACCATT	660
OY	661	GGAACAAACACAGCAGAAATCAGATTTTCCAGCTCTTAGAGATTTGCCCGCTGGAATG	720
Db	661	GGAACAAACACAGCAGAAATCAGATTTTCCAGCTCTTAGAGATTTGCCCGCTGGAATG	720
OY	721	TATGAAATCCGGTTGACCCGCGCAAGGACAGGGAAGGCAAGAAATCAATCTGAGCGTT	780
Db	721	TATGAAATCCGGTTGACCCGCGCAAGGACAGGGAAGGCAAGAAATCAATCTGAGCGTT	780
OY	781	GCCAAACAGGGAATTTAGTGTTCAGGGTTTCACTAAGATCAATGCCCTCAACTGAGCC	840
Db	781	GCCAAACAGGGAATTTAGTGTTCAGGGTTTCACTAAGATCAATGCCCTCAACTGAGCC	840
OY	841	AAAGTGCAGAACTGAGCTTCAAGAGGAAGGCTTTCTCAATCAAGCTCCGCGCAGATGCC	900
Db	841	AAAGTGCAGAACTGAGCTTCAAGAGGAAGGCTTTCTCAATCAAGCTCCGCGCAGATGCC	900
OY	901	AATAGTGGTATCAGAGATACCTTGAATTTCTGATGCGCACTCGGGAATTTCTGCAATGC	960
Db	901	AATAGTGGTATCAGAGATACCTTGAATTTCTGATGCGCACTCGGGAATTTCTGCAATGC	960
OY	961	TTCTGGAATAATCTGTGTGAACATCAGCTCTTTTGAAGCTTTTGAAGGCCCAACCA	1020

Db	961	TTCTGGAAAAATCTGTGTTGAACATCATGCTTCTTTTGAATTTTGAAGGCCAAACCA	1020
QY	1021	AAGCCCAAGCCCGCTCTTTAAGCCGGGGGTCAATCATTTTCGTTCACTGTGCGACTCAG	1080
Db	1021	AAGCCCAAGCCCGCTCTTTAAGCCGGGGGTCAATCATTTTCGTTCACTGTGCGACTCAG	1080
QY	1081	AAGCAGGTTCTCGACTATGTTTAAGAAGAGGACATTAAGAAAGTGCAGTTTGAAGAGAG	1140
Db	1081	AAGCAGGTTCTCGACTATGTTTAAGAAGAGGACATTAAGAAAGTGCAGTTTGAAGAGAG	1140
QY	1141	CACAGCAAGATTCATCTATCCGAGGCTTGCTCAACAGCTACAGAACTGAATTCGGAA	1200
Db	1141	CACAGCAAGATTCATCTATCCGAGGCTTGCTCAACAGCTACAGAACTGAATTCGGAA	1200
QY	1201	GTGCTGGAGCAGTCTCAGCAGAGCACCAGCTTACATTTTGAAGAGGTGCGAATCTCCA	1260
Db	1201	GTGCTGGAGCAGTCTCAGCAGAGCACCAGCTTACATTTTGAAGAGGTGCGAATCTCCA	1260
QY	1261	GGGGGCCCAAGCTGTCGGCCGAGAAAGAAACCGAAGGTTTCCGCGGGGAGGCGGGGTG	1320
Db	1261	GGGGGCCCAAGCTGTCGGCCGAGAAAGAAACCGAAGGTTTCCGCGGGGAGGCGGGGTG	1320
QY	1321	CACCCGAGCCCTGCGCCGAGAGAGAGCCCGCGGGTTACAAAGCAGGCGGACCGAGCCGC	1380
Db	1321	CACCCGAGCCCTGCGCCGAGAGAGAGCCCGCGGGTTACAAAGCAGGCGGACCGAGCCGC	1380
QY	1381	TCGGCGGCCCAACGAGAGAAAGAGAGAGTGTTAAGAGTAAAGCCAGCAGAGTAAACCT	1440
Db	1381	TCGGCGGCCCAACGAGAGAAAGAGAGAGTGTTAAGAGTAAAGCCAGCAGAGTAAACCT	1440
QY	1441	CAGCCCCCGCAGCCAGACACAGGCTCCCTGACCTGAGCTCTCACTTTTCCGAGCTGCT	1500
Db	1441	CAGCCCCCGCAGCCAGACACAGGCTCCCTGACCTGAGCTCTCACTTTTCCGAGCTGCT	1500
QY	1501	GTGAACCTCGAGGGGGGAGTGGCCCGTGCACAGTGAACCTTGTCTCCCAACTGAGGCC	1560
Db	1501	GTGAACCTCGAGGGGGGAGTGGCCCGTGCACAGTGAACCTTGTCTCCCAACTGAGGCC	1560
QY	1561	GACACCAAGCAGGCTCTCCCTTGATCAGCCGCTGTGTAATGACCAAGGCTGCCCCGG	1620
Db	1561	GACACCAAGCAGGCTCTCCCTTGATCAGCCGCTGTGTAATGACCAAGGCTGCCCCGG	1620
QY	1621	ACGACGATGAGTGAAGGCGCGGAGGAAAGATTCCTCAACTGATTAAGGCTACTTCATA	1680
Db	1621	ACGACGATGAGTGAAGGCGCGGAGGAAAGATTCCTCAACTGATTAAGGCTACTTCATA	1680
QY	1681	GCTAAAGAAAGTGTCTACACCGAGCGAAACATATCTGAAGGATCTCGAAAGTTATCACTTCG	1740
Db	1681	GCTAAAGAAAGTGTCTACACCGAGCGAAACATATCTGAAGGATCTCGAAAGTTATCACTTCG	1740
QY	1741	TGCTTTGAGAGCAGTGAAGCAAAAGAGAGCGCATGCGGAGAGCACTGAAGAAAGTTCATA	1800
Db	1741	TGCTTTGAGAGCAGTGAAGCAAAAGAGAGCGCATGCGGAGAGCACTGAAGAAAGTTCATA	1800
QY	1801	TTCCCGAATTTTGAACCTTTGACCAAAATTTTCAATCTAATTTTCTCAAGGAAATTTGACAA	1860
Db	1801	TTCCCGAATTTTGAACCTTTGACCAAAATTTTCAATCTAATTTTCTCAAGGAAATTTGACAA	1860
QY	1861	CGAATTGCCCCCTGTGGGAAAGCCGCTCAAAATGCCCCAAATCAGAGATTACAAAGAGATGGC	1920
Db	1861	CGAATTGCCCCCTGTGGGAAAGCCGCTCAAAATGCCCCAAATCAGAGATTACAAAGAGATGGC	1920
QY	1921	GATGTCATGCTGAAGAACTTCAAGGGCATGAAGCACCCTGCGGCTCACTGTGGAACAC	1980
Db	1921	GATGTCATGCTGAAGAACTTCAAGGGCATGAAGCACCCTGCGGCTCACTGTGGAACAC	1980
QY	1981	AGCGAGGCTTTGAGAGGCTTGAAGAAATGAATCAAGAGCTCCCGCGGCTGAGAACTTC	2040
Db	1981	AGCGAGGCTTTGAGAGGCTTGAAGAAATGAATCAAGAGCTCCCGCGGCTGAGAACTTC	2040
QY	2041	TGCAGAGACTTTGAGCTGCAAGAGGTGTATTACTACCGTCAACACTTCTCTCTGCGG	2100

```

Db      2041 TGCAGAGACTTTGAGCTGCAGAAAGTGTTACTTACCGCTCAACACTTCTCTGCGG 2100
Qy      2101 CCAGTGCACCGGCTGATGCACTACAGCAGAGTCTTGAGAGGGCTGTGCAAAACCAACCGG 2160
Db      2101 CCAGTGCACCGGCTGATGCACTACAGCAGAGTCTTGAGAGGGCTGTGCAAAACCAACCGG 2160
Qy      2161 CCGAGCAGCGCCGACTTCAGGGAATGCGGAGCCGCTTTGGCAGAGATCAACGAGATGATG 2220
Db      2161 CCGAGCAGCGCCGACTTCAGGGAATGCGGAGCCGCTTTGGCAGAGATCAACGAGATGATG 2220
Qy      2221 GCACAGCTCCACCGGATCGATGATCAAGATGGAATTTCCAGAAAGCTGCAGCACTCAAG 2280
Db      2221 GCACAGCTCCACCGGATCGATGATCAAGATGGAATTTCCAGAAAGCTGCAGCACTCAAG 2280
Qy      2281 AAAAGATTTGATTTGGCATTTGATCAATCTTGTGTTCCGGGAAAGGAGTTCACTCGTGGGC 2340
Db      2281 AAAAGATTTGATTTGGCATTTGATCAATCTTGTGTTCCGGGAAAGGAGTTCACTCGTGGGC 2340
Qy      2341 AGCCTCAGCAGAGCTCCGGGGAAGGGGCTCCAGCAGGCACTGTTCTTCCGTGCAAGCAG 2400
Db      2341 AGCCTCAGCAGAGCTCCGGGGAAGGGGCTCCAGCAGGCACTGTTCTTCCGTGCAAGCAG 2400
Qy      2401 GTCTGCTATATACAGAGCCGGGGCTGACGGCTCCCAATCAGTTTAAAGTCCACGGGCAAG 2460
Db      2401 GTCTGCTATATACAGAGCCGGGGCTGACGGCTCCCAATCAGTTTAAAGTCCACGGGCAAG 2460
Qy      2461 CTCCGCTCTATGAGCATGACGATTTGAGAGAGGAAAGCAGATGAGGGGGTGCCTCCACTGC 2520
Db      2461 CTCCGCTCTATGAGCATGACGATTTGAGAGAGGAAAGCAGATGAGGGGGTGCCTCCACTGC 2520
Qy      2521 CTGACCTCCGGGGCCAGGGGCGAGTCCATCATCTGTGGCCGCGAGTTCTCGGTCCGAGATG 2580
Db      2521 CTGACCTCCGGGGCCAGGGGCGAGTCCATCATCTGTGGCCGCGAGTTCTCGGTCCGAGATG 2580
Qy      2581 GAGAACTGGGTTGAGGACATCCAGATGGCCATTGACCTGCGGGAAGAGCAGCAGAGCCCC 2640
Db      2581 GAGAACTGGGTTGAGGACATCCAGATGGCCATTGACCTGCGGGAAGAGCAGCAGAGCCCC 2640
Qy      2641 GCGCCCTGAGTTCTTGCGCCAGCAGCCCTCTGACAAACAAGTCCCTCTATGAAAGCAGCCGG 2700
Db      2641 GCGCCCTGAGTTCTTGCGCCAGCAGCCCTCTGACAAACAAGTCCCTCTATGAAAGCAGCCGG 2700
Qy      2701 GCTGACCAAGAGTCAGAGATGACCTGAGCGCTCCGCGACATCGCTGAGCGCCAGGCC 2760
Db      2701 GCTGACCAAGAGTCAGAGATGACCTGAGCGCTCCGCGACATCGCTGAGCGCCAGGCC 2760
Qy      2761 CCGCACCGCGGCAACAATGTGTGCACTGTGTGCGACCGCAACAACCGCTCTCATG 2820
Db      2761 CCGCACCGCGGCAACAATGTGTGCACTGTGTGCGACCGCAACAACCGCTCTCATG 2820
Qy      2821 GTGGAATTCAGATCGCAGTGCAGAAATCACTGTTCTGGAACCTGCTGAGGAATTCAAA 2880
Db      2821 GTGGAATTCAGATCGCAGTGCAGAAATCACTGTTCTGGAACCTGCTGAGGAATTCAAA 2880
Qy      2881 AACAGCAAGGGGTGGGAGAAAGTGGGTGTGTTCAAAACCTTCGCTGTTCTTCAAC 2940
Db      2881 AACAGCAAGGGGTGGGAGAAAGTGGGTGTGTTCAAAACCTTCGCTGTTCTTCAAC 2940
Qy      2941 AAATCACAACCAAGACAAATCATCCCTTGGCAGCCTCTCTGCTCGGCTACTGCTCACC 3000
Db      2941 AAATCACAACCAAGACAAATCATCCCTTGGCAGCCTCTCTGCTCGGCTACTGCTCACC 3000
Qy      3001 ATCCCTCTGAGTCCGAGAAACATCCGAAAGACTACGTGTTCAAGCTGCACCTTCAAGTCC 3060
Db      3001 ATCCCTCTGAGTCCGAGAAACATCCGAAAGACTACGTGTTCAAGCTGCACCTTCAAGTCC 3060
Qy      3061 CACGTCTACTACTTCAGGCGGAGAAAGCAGATACAGTTGCAAAAGTGGATGGAAGTATC 3120
Db      3061 CACGTCTACTACTTCAGGCGGAGAAAGCAGATACAGTTGCAAAAGTGGATGGAAGTATC 3120
Qy      3121 CCGAGTGCACCAAGCTCTGCTCTGCGAACCACAGTGTGAGCCACAAGAAGTCTTGTG 3180
Db      3121 CCGAGTGCACCAAGCTCTGCTCTGCGAACCACAGTGTGAGCCACAAGAAGTCTTGTG 3180

```

```

Qy      3181 TATTGATGGCGGAGACACTCGTTCGAGTGGGCTTTCTTCTGSAAGAGCTTCTT 3240
Db      3181 TATTGATGGCGGAGACACTCGTTCGAGTGGGCTTTCTTCTGSAAGAGCTTCTT 3240
Qy      3241 CTTCTGTATTAAATGAAAGCTGTGTAATAATCAACCTGTCTGAAATCAAAAAATGAGCTT 3300
Db      3241 CTTCTGTATTAAATGAAAGCTGTGTAATAATCAACCTGTCTGAAANCAAAAAATGAGCTT 3300
Qy      3301 CCGAGCAGCTCTCTGCTCTCTCACAAGCCGCGGTTTTTTTAAACCCGAGCTCTCAGGTTGAA 3360
Db      3301 CCGAGCAGCTCTCTGCTCTCTCACAAGCCGCGGTTTTTTTAAACCCGAGCTCTCAGGTTGAA 3360
Qy      3361 TGAACAGCGGCTCCACTCCATCCCTGAGATCCGCTGGGGGGGCTTCTTACGTAAG 3420
Db      3361 TGAACAGCGGCTCCACTCCATCCCTGAGATCCGCTGGGGGGGCTTCTTACGTAAG 3420
Qy      3421 CCAATATTAAACATTTGCAAT 3442
Db      3421 CCAATATTAAACATTTGCAAT 3442

```

```

RESULT 2
AEB03370
ID      AEB03370 standard; DNA; 3442 BP.
XX
AC      AEB03370;
XX
DT      09-FEB-2006 (first entry)
XX
DE      Acute myelogenous leukemia prognosis related DNA sequence SEQ ID NO: 31.
XX
KW      ds; gene; acute myelogenous leukemia; prognosis; gene expression;
XX      biochip.
OS      Homo sapiens.
XX
PN      JP200533987-A.
XX
PD      08-DEC-2005.
XX
PF      06-MAY-2005; 2005JP-00135284.
XX
PR      06-MAY-2004; 2004US-0566635P.
XX
PA      (VERI-) VERIDEX LLC.
XX
PI      Laponi M;
XX
PS      WPI; 2006-004067/01.
XX
PT      Evaluating an acute myelogenous leukemia patient's prognosis, comprising
PT      detecting lower or higher expression level of gene recognized by probe
PT      set e.g. 202820-at and 206148-at, with respect to predetermined cut-off
PT      level.
XX
PS      Example 5; SEQ ID NO 31; 60pp; Japanese.
XX
CC      The present sequence is that of a human gene sequence which is claimed
CC      for use in evaluating the prognosis of patients suffering from acute
CC      myelogenous leukemia (AML) by analysis of the patients gene expression
CC      profile. The invention relates to a novel method for evaluating the
CC      prognosis of a patient with acute myelogenous leukemia by detecting
CC      higher/lower expression level of genes encoding mRNA recognized by a
CC      probe set chosen from 19 nucleotide sequences (AEB03340 or AEB03344-
CC      AEB03361). Also, gene expression profiling of AML patients using a set of
CC      probes based on 167 sequences (AEB03362-AEB03527) using a biochip was
CC      performed to determine prognosis. Methods are also included for
CC      determining an AML patient's state, determining AML patient's treatment,
CC      protocol, determining whether the patient will respond to the treatment,
CC      and producing an AML patient's prognosis report by analysis of the
CC      patients gene expression profile. The methods are useful in evaluating an
CC      AML patient's prognosis and for treating AML patient.

```

XX Sequence 3442 BP; 864 A; 952 C; 927 G; 699 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 3442; DB 15; Length 3442;
Beet Local Similarity 100.0%; Pred. No. 0;
Matches 3442; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 CCCCCAGCCGCGCGCTGTGGAGATATTTCTTAAGCCGCTTTCATCATGGAGAAATA 60
1 CCCCCAGCCGCGCGCGCTGTGGAGATATTTCTTAAGCCGCTTTCATCATGGAGAAATA 60
61 GAGCAGAGGCCGACCCCAAGATACAGATCGGGGGCCCCGGAAATTCGGGATCACTAC 120
61 GAGCAGAGGCCGACCCCAAGATACAGATCGGGGGCCCCGGAAATTCGGGATCACTAC 120
121 TTGGAACGTGGAAGAGAGCCGCCCCCAACCTTCAGGAAACTCGTGTCCATCAAAATC 180
121 TTGGAACGTGGAAGAGAGCCGCCCCCAACCTTCAGGAAACTCGTGTCCATCAAAATC 180
181 CAGATGCTGGATGACACCCAGAGAGGCAATTTGAAGTTCCAAAGAGCTCTGGGAAGTG 240
181 CAGATGCTGGATGACACCCAGAGAGGCAATTTGAAGTTCCAAAGAGCTCTGGGAAGTG 240
241 CTGCTGGATGCAAGTTTGCAACCACTCAACCTGTGGAAGGTGACTATTTTGCCCTCGAG 300
241 CTGCTGGATGCAAGTTTGCAACCACTCAACCTGTGGAAGGTGACTATTTTGCCCTCGAG 300
301 TTTCCTGATCAGAAAAGATCACGGTGTGGCTGGAATCTCTTAAACCCATTTGGAACG 360
301 TTTCCTGATCAGAAAAGATCACGGTGTGGCTGGAATCTCTTAAACCCATTTGGAACG 360
361 ATTAGAGGCCAAAGCAGCTGTGTTTGAAGTTGTGTGAAATCTTTTCCGCTGACAC 420
361 ATTAGAGGCCAAAGCAGCTGTGTTTGAAGTTGTGTGAAATCTTTTCCGCTGACAC 420
421 ACACACTCTCAAGAAATCTCACAGGTACCTGTTGCGCTGAGTGAAGCAGACCTTG 480
421 ACACACTCTCAAGAAATCTCACAGGTACCTGTTGCGCTGAGTGAAGCAGACCTTG 480
481 GGTCAAGGCGAGGTGACGTGTAATGACACAGGCGAGCTCTCTGATTTTCAACATTTG 540
481 GGTCAAGGCGAGGTGACGTGTAATGACACAGGCGAGCTCTCTGATTTTCAACATTTG 540
541 CAATCGAGATTGGGATTTTGAATGAGAGCTTGACAGAGACACTTACCAAAAATATA 600
541 CAATCGAGATTGGGATTTTGAATGAGAGCTTGACAGAGACACTTACCAAAAATATA 600
601 TACATTAATCTCAGCAAGACGCACTAGAGGACAAATCTGTGAAATTCACCATTAACA 660
601 TACATTAATCTCAGCAAGACGCACTAGAGGACAAATCTGTGAAATTCACCATTAACA 660
661 GGACAAACACGACGAGATCAGATTTTCAGCTCTTAGAGATTGCCCGTGGCTAGAGATG 720
661 GGACAAACACGACGAGATCAGATTTTCAGCTCTTAGAGATTGCCCGTGGCTAGAGATG 720
721 TATGGAATCCGTTGACCCGCGCAAGGACAGGAAAGGCAAGAAATCAATCTGGCCGTT 780
721 TATGGAATCCGTTGACCCGCGCAAGGACAGGAAAGGCAAGAAATCAATCTGGCCGTT 780
781 GCGAACACGGGAATTTAGTGTTCAGGGTTTCACTAGATCAATGCTTCAACTGGGGC 840
781 GCGAACACGGGAATTTAGTGTTCAGGGTTTCACTAGATCAATGCTTCAACTGGGGC 840
841 AAGGTGCGGAAGCTGAGTTCAAGAGAGCGCTTTCATCAAGCTCCGGCCAGATGCC 900
841 AAGGTGCGGAAGCTGAGTTCAAGAGAGCGCTTTCATCAAGCTCCGGCCAGATGCC 900
901 AATAGTGCCTTACAGAGATCTTTGAATTCCTGATGCGCAGTCGGGATTTCTGCAAGTCC 960
901 AATAGTGCCTTACAGAGATCTTTGAATTCCTGATGCGCAGTCGGGATTTCTGCAAGTCC 960
961 TTCTGGAAATCTGTGTTGAACATCATGCTTCTTTAGCTTTTGAAGGCCCAACCA 1020
961 TTCTGGAAATCTGTGTTGAACATCATGCTTCTTTAGCTTTTGAAGGCCCAACCA 1020

1021 TTCTGGAAATCTGTGTTGAACATCATGCTTCTTTAGCTTTTGAAGGCCCAACCA 1020
1021 TTCTGGAAATCTGTGTTGAACATCATGCTTCTTTAGCTTTTGAAGGCCCAACCA 1020
1021 AAGCCCAAGCCGCTCCTTTAGCCGGGGATCATCATTTTGGTTCACTGTCGACTGAG 1080
1021 AAGCCCAAGCCGCTCCTTTAGCCGGGGATCATCATTTTGGTTCACTGTCGACTGAG 1080
1081 AAGCAGTTTCTGACTATGTTAAAGAGAGACATTAAGAGTGCAGTTTGAAGGAAG 1140
1081 AAGCAGTTTCTGACTATGTTAAAGAGAGACATTAAGAGTGCAGTTTGAAGGAAG 1140
1141 CACAGCAAGATTCATTCATCCGAGCCCTTGCTTACAGCCCTACAGAACTGAATTCGAA 1200
1141 CACAGCAAGATTCATTCATCCGAGCCCTTGCTTACAGCCCTACAGAACTGAATTCGAA 1200
1201 GTGCTGAGACAGTCTCAGCAGAGCACAGCCCTTACATTTTGAAGAGTGCAGAACTCCA 1260
1201 GTGCTGAGACAGTCTCAGCAGAGCACAGCCCTTACATTTTGAAGAGTGCAGAACTCCA 1260
1261 GGGGGCCAGAGCTGCGCGGAGAAAGAACGAAAGTTTCCGCGGGAGCCGGGGTTCG 1320
1261 GGGGGCCAGAGCTGCGCGGAGAAAGAACGAAAGTTTCCGCGGGAGCCGGGGTTCG 1320
1321 CACCCGAGCCCTGCGCGGAGAGAAAGCCCGCGGTAAACAAGACGGCGAGACCCGCG 1380
1321 CACCCGAGCCCTGCGCGGAGAGAAAGCCCGCGGTAAACAAGACGGCGAGACCCGCG 1380
1381 TCGGCGCCACGAGAGAGAGAGAGGTGTTAAGATTAAGATCCACAGAGTAAACT 1440
1381 TCGGCGCCACGAGAGAGAGAGAGGTGTTAAGATTAAGATCCACAGAGTAAACT 1440
1441 CAGCCCCGACAGCCAGAGCAAGGCTCCCTGACTGGCAGTCTCACCTTTCCGAGCTGCT 1500
1441 CAGCCCCGACAGCCAGAGCAAGGCTCCCTGACTGGCAGTCTCACCTTTCCGAGCTGCT 1500
1501 GTGAACCTGACAGGGGGAGTGCGCCCTGCAACGTAACCTTGTCTCCAACTGAGCCCC 1560
1501 GTGAACCTGACAGGGGGAGTGCGCCCTGCAACGTAACCTTGTCTCCAACTGAGCCCC 1560
1561 GACACCAAGCAGGCTCTCCCTTGATCAGCCGCTGCTGTAATGACAGAGCCTGCCCGG 1620
1561 GACACCAAGCAGGCTCTCCCTTGATCAGCCGCTGCTGTAATGACAGAGCCTGCCCGG 1620
1621 ACGAGCATGAGATGAGAGGCGCGAGAGAAAGATTCCTCACTGATTAAGCTTACATA 1680
1621 ACGAGCATGAGATGAGAGGCGCGAGAGAAAGATTCCTCACTGATTAAGCTTACATA 1680
1681 GCTAAGAGAGTGTCTACACCGAGGGAATATCTGAAGAGATCTGGAAGTTATCACTTCG 1740
1681 GCTAAGAGAGTGTCTACACCGAGGGAATATCTGAAGAGATCTGGAAGTTATCACTTCG 1740
1741 TGGTTTCAGAGCAGTAGAGCAAGAGAGCCATGCGGAAAGCACTGAAAGTCTCATA 1800
1741 TGGTTTCAGAGCAGTAGAGCAAGAGAGCCATGCGGAAAGCACTGAAAGTCTCATA 1800
1801 TTCCGGAATTTTGAACCTTTGACAAATTTCACTAATTTTCTCAAGAAATTAAGCAA 1860
1801 TTCCGGAATTTTGAACCTTTGACAAATTTCACTAATTTTCTCAAGAAATTAAGCAA 1860
1861 TTCCGGAATTTTGAACCTTTGACAAATTTCACTAATTTTCTCAAGAAATTAAGCAA 1860
1861 TTCCGGAATTTTGAACCTTTGACAAATTTCACTAATTTTCTCAAGAAATTAAGCAA 1860
1861 GACTTGGCCCTGTGGAGAGGCGCTCAATGCGCCAAATCAAGATTAACAAAGAAATCGGG 1920
1861 GACTTGGCCCTGTGGAGAGGCGCTCAATGCGCCAAATCAAGATTAACAAAGAAATCGGG 1920
1921 GATGTCATGCTGAAGAAATTCAGGGCATGAAGCACTGCGGCTCCTGCTGGAAGCAC 1980
1921 GATGTCATGCTGAAGAAATTCAGGGCATGAAGCACTGCGGCTCCTGCTGGAAGCAC 1980
1981 AAGCAGGCTTTGAGAGGCTTGAAGATGAAGATCAAGAGACTCCCGCGGCTGGAAGAACTTC 2040
1981 AAGCAGGCTTTGAGAGGCTTGAAGATGAAGATCAAGAGACTCCCGCGGCTGGAAGAACTTC 2040
2041 TGCAGAGACTTTGAGCTGCAAGAGGTGTTAACCTACCGCTCAACACTTCTCTCGCGG 2100
2041 TGCAGAGACTTTGAGCTGCAAGAGGTGTTAACCTACCGCTCAACACTTCTCTCGCGG 2100

2101 CCAGTGCACCGGCTCATGCACTACAGAGCTCTGAGAGCGGCTGCAAAACACACCGG 2160
2101 CCAGTGCACCGGCTCATGCACTACAGAGCTCTGAGAGCGGCTGCAAAACACACCGG 2160
2161 CCGAGCCAGCGGCACTTCAGAGGAGCTGCGAGCGCTTTGGCAGAGATCAACGAGATGATG 2220
2161 CCGAGCCAGCGGCACTTCAGAGGAGCTGCGAGCGGCTTTGGCAGAGATCAACGAGATGATG 2220
2221 GCACAGCTCCAGGATTCAGATGATCAAGATGAGAAATTTCCAGAGCTGCAACGATCAAG 2280
2221 GCACAGCTCCAGGATTCAGATGATCAAGATGAGAAATTTCCAGAGCTGCAACGATCAAG 2280
2281 AAGAGATTGATGGGATTCAGATCTTGTGTCTCCGGGAAAGGAGTTCAATCGCTGGGAG 2340
2281 AAGAGATTGATGGGATTCAGATCTTGTGTCTCCGGGAAAGGAGTTCAATCGCTGGGAG 2340
2341 AGCCTCAGCAAGCTCTCGGGGAAAGGGGCTCAGCAGCGCATGTTCTTCTGTTCAACGAC 2400
2341 AGCCTCAGCAAGCTCTCGGGGAAAGGGGCTCAGCAGCGCATGTTCTTCTGTTCAACGAC 2400
2401 GTCTGCTATATACAGAGCCGGGGGCTGACGGGCTTCAATCAATTTAAAGTCCAGGGCAG 2460
2401 GTCTGCTATATACAGAGCCGGGGGCTGACGGGCTTCAATCAATTTAAAGTCCAGGGCAG 2460
2461 CTCCGGCTTATGAGTGAAGATGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2520
2461 CTCCGGCTTATGAGTGAAGATGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2520
2521 CTGACCTCCGGGGGCGAGCGGCGAGTCCATCATCTGTGGCGCGCATCTCGGATCCAGATG 2580
2521 CTGACCTCCGGGGGCGAGCGGCGAGTCCATCATCTGTGGCGCGCATCTCGGATCCAGATG 2580
2581 GAGAGTGGGATGAGGACATCCAGATGCGCATTTGACCTGCGGAGGAGGAGGAGGAGGAGG 2640
2581 GAGAGTGGGATGAGGACATCCAGATGCGCATTTGACCTGCGGAGGAGGAGGAGGAGGAGG 2640
2641 GCGCCCTGAGTTCTGCGCCAGCAGCCCCCTGACAAACAGTCCCTGATGAGGAGGAGGAGG 2700
2641 GCGCCCTGAGTTCTGCGCCAGCAGCCCCCTGACAAACAGTCCCTGATGAGGAGGAGGAGG 2700
2701 GGTGACCAAGAGTCAAGAGATGACCTGAGCGGCTCGCGCAATCGTGAAGCGCCAGGCG 2760
2701 GGTGACCAAGAGTCAAGAGATGACCTGAGCGGCTCGCGCAATCGTGAAGCGCCAGGCG 2760
2761 CCGCAGCGGCGGCAACAAATGATGACGATGCTGCGCAGCAGCAACGAGCGCTTCATG 2820
2761 CCGCAGCGGCGGCAACAAATGATGACGATGCTGCGCAGCAGCAACGAGCGCTTCATG 2820
2821 GTGAGCTTACGATCGCAGTGGAGATCACTGTTGTCTGAGAAACCTGCTGAGAAATTCAAA 2880
2821 GTGAGCTTACGATCGCAGTGGAGATCACTGTTGTCTGAGAAACCTGCTGAGAAATTCAAA 2880
2881 AACAGCAACGGGAGGAGAGCTGAGGATGTTTCAAACTTGTGCGCTGTTCTTCTAC 2940
2881 AACAGCAACGGGAGGAGAGCTGAGGATGTTTCAAACTTGTGCGCTGTTCTTCTAC 2940
2881 AACAGCAACGGGAGGAGAGCTGAGGATGTTTCAAACTTGTGCGCTGTTCTTCTAC 2940
2941 AAATCAACACGAGCAATCATCTCCCTTGCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3000
2941 AAATCAACACGAGCAATCATCTCCCTTGCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3000
3001 ATCCCTCTGAGTCCAGAAATCCAGAAAGATTAAGTGTTCAGAGCTTCAAGAGTCC 3060
3001 ATCCCTCTGAGTCCAGAAATCCAGAAAGATTAAGTGTTCAGAGCTTCAAGAGTCC 3060
3061 CAGCTTACTACTTCAAGGCGGAAACGAGTACAGCTTCAAAAGGATGAGTGAAGTATC 3120
3061 CAGCTTACTACTTCAAGGCGGAAACGAGTACAGCTTCAAAAGGATGAGTGAAGTATC 3120
3121 CCGAGTGCACCAAGCTCTGCTGCGAACCACGATGTTGAGCCAAAGAGTCTTTGAG 3180
3121 CCGAGTGCACCAAGCTCTGCTGCGAACCACGATGTTGAGCCAAAGAGTCTTTGAG 3180
3121 CCGAGTGCACCAAGCTCTGCTGCGAACCACGATGTTGAGCCAAAGAGTCTTTGAG 3180

3181 TATTGATGGCGGAGACACTGTTCCGAGTGGGCTGCTTCTGGAAGACGTTCTTT 3240
3181 TATTGATGGCGGAGACACTGTTCCGAGTGGGCTGCTTCTGGAAGACGTTCTTT 3240
3241 CTTCGTATTAATGAAAGCTGTGTAATTAACACCTGTCTGAAATCAAAAATGAGCTT 3300
3241 CTTCGTATTAATGAAAGCTGTGTAATTAACACCTGTCTGAAATCAAAAATGAGCTT 3300
3301 CCCAGAGCTCTCTGCTGCTGCAAGCGGCTTTTAAACCCGAGCCTCTCAGCGTTTGA 3360
3301 CCCAGAGCTCTCTGCTGCTGCAAGCGGCTTTTAAACCCGAGCCTCTCAGCGTTTGA 3360
3361 TGAACAGCGCTCCCACTCCAGCTGCGATCGCTGAGGAGGAGGAGGAGGAGGAGGAGG 3420
3361 TGAACAGCGCTCCCACTCCAGCTGCGATCGCTGAGGAGGAGGAGGAGGAGGAGGAGG 3420
3421 CCAATTTAAACATTTGCTATT 3442
3421 CCAATTTAAACATTTGCTATT 3442

RESULT 3

AE22714
ID AEB22714 standard; DNA; 5028 BP.

AC AEB22714;

DT 22-SEP-2005 (first entry)

DE Human colon cancer-associated gene SEQ ID NO:9.

KW cancer; genetic marker; lymphoma; cytostatic; hematological disease;

KW immune disorder; neoplasm; colon tumor; gastrointestinal disease;

KW adenocarcinoma; carcinoma; screening; diagnostic; prognostic;

KW pharmaceutical; ds; gene.

OS Homo sapiens.

PN WO2005064009-A1.

PD 14-JUL-2005.

PF 23-DEC-2004; 2004WO-DK000914.

PR 27-DEC-2003; 2003DK-0001940.

PR 24-JAN-2004; 2004DK-0000096.

PR 07-APR-2004; 2004DK-0000586.

PR 26-NOV-2004; 2004DK-00001843.

PA (AROS-) AROS APPLIED BIOTECHNOLOGY APS.

PI Ornofit TF, Jensen JL, Kruhoffer M, Laiho P, Aaltonen LA;

PT WPI; 2005-506661/51.

PS Claim 9; SEQ ID NO 9; 338bp; English.

CC The invention relates to a method of classifying the cancer in an

CC individual having contracted cancer, which comprises determining the

CC microsatellite status of the tumor. The microsatellite status or the

CC hereditary or sporadic nature of the cancer is determined by a method

CC comprising: (a) obtaining a sample from the individual having contracted

CC cancer, the sample comprising gene expression products; the presence

CC and/or amount of which forms a pattern that is indicative of the

CC microsatellite status or the hereditary or sporadic nature of the

CC products forming the pattern, and obtaining an indication of the

CC microsatellite status or the hereditary or sporadic nature of the

CC expression products forming a pattern used to determine the

QY 1801 TTCCCGAATTTTGAACCTTTGCACAAATTTTCTAATTTTCTCAAGGAAATTGAGCAA 1860
 DB 2089 TTCCCGAATTTTGAACCTTTGCACAAATTTTCTAATTTTCTCAAGGAAATTGAGCAA 2148
 QY 1861 CGACTTCCCTGTGGAGAGCCGCTCAATGCCCAATCAGAGATTACCAAAAGATCGGC 1920
 DB 2149 CGACTTCCCTGTGGAGAGCCGCTCAATGCCCAATCAGAGATTACCAAAAGATCGGC 2208
 QY 1921 GATGTATGTGAGAAACATTCAGGGGATGAGACCTTGGCGGCTCACTGTGGAAGCAC 1980
 DB 2209 GATGTATGTGAGAAACATTCAGGGGATGAGACCTTGGCGGCTCACTGTGGAAGCAC 2268
 QY 1981 AGCGAGGCGCTTGGAGGCGCTTGGAGATGGAATCAAGATCCCGGGGCGCTGGAGACTTC 2040
 DB 2269 AGCGAGGCGCTTGGAGGCGCTTGGAGATGGAATCAAGATCCCGGGGCGCTGGAGACTTC 2328
 QY 2041 TGCAGAGACTTTGAGCTGAGAGAGGTGTACTTACCTGCTCAACACTTCTCTGCGG 2100
 DB 2329 TGCAGAGACTTTGAGCTGAGAGAGGTGTACTTACCTGCTCAACACTTCTCTGCGG 2388
 QY 2101 CCACTGACCGGCTCATGCACTACAGACAGCTCTTGAAGCGCTGTGCAACACACCGG 2160
 DB 2389 CCACTGACCGGCTCATGCACTACAGACAGCTCTTGAAGCGCTGTGCAACACACCGG 2448
 QY 2161 CGAGGCGGCGGACTTCAAGGGAATGCGAGCGGCTTTGGCAGAGATCAAGAGATGATG 2220
 DB 2449 CGAGGCGGCGGACTTCAAGGGAATGCGAGCGGCTTTGGCAGAGATCAAGAGATGATG 2508
 QY 2221 GCACAGCTCCAGCGTACGATGATCAAGATGAGAGAAATTTCCAGAGCTGACGAACTCAAG 2280
 DB 2509 GCACAGCTCCAGCGTACGATGATCAAGATGAGAGAAATTTCCAGAGCTGACGAACTCAAG 2568
 QY 2281 AAAAGTTTATTTGGCATTTGACAACTTTGTGTTCCGGGAAAGGAGTTCACTCGTCTGGG 2340
 DB 2569 AAAAGTTTATTTGGCATTTGACAACTTTGTGTTCCGGGAAAGGAGTTCACTCGTCTGGG 2628
 QY 2341 AGCTCAGAGAGCTCCGGGGGAAAGGGGCTCCAGCAGAGATGTTCTTCCGTTCAACGAGC 2400
 DB 2629 AGCTCAGAGAGCTCCGGGGGAAAGGGGCTCCAGCAGAGATGTTCTTCCGTTCAACGAGC 2688
 QY 2401 GTCTGCTATTAACGAGCGCGGGGCTGAGCGGCTCCAAATCAAGTTTAAAGTCCACGGGCGAG 2460
 DB 2689 GTCTGCTATTAACGAGCGCGGGGCTGAGCGGCTCCAAATCAAGTTTAAAGTCCACGGGCGAG 2748
 QY 2461 GTCCCGCTCTATGAGCATGACGATTGAGAGAGCGAAAGCAGAGTGGGGGGTCCCACTGC 2520
 DB 2749 GTCCCGCTCTATGAGCATGACGATTGAGAGAGCGAAAGCAGAGTGGGGGGTCCCACTGC 2808
 QY 2521 CTGACCTCCGGGGGCGAGGGGCGAGTCAATCATGTTGCGGCGGCTTCTGCTCCGAGATG 2580
 DB 2809 CTGACCTCCGGGGGCGAGGGGCGAGTCAATCATGTTGCGGCGGCTTCTGCTCCGAGATG 2868
 QY 2581 GAGAAAGTGGGTTGAGAGCATCCAGATGCGCATTTGACTTGGCGGAGAAAGACAGCAGCCCC 2640
 DB 2869 GAGAAAGTGGGTTGAGAGCATCCAGATGCGCATTTGACTTGGCGGAGAAAGACAGCAGCCCC 2928
 QY 2641 GCGCCCTGAGTTCTTGCGCAGACGCCCCCTGTACACAAGTCCCTGATGAAGCCACCGCG 2700
 DB 2929 GCGCCCTGAGTTCTTGCGCAGACGCCCCCTGTACACAAGTCCCTGATGAAGCCACCGCG 2988
 QY 2701 GGTGACCAAGAGTCAAGAGATGACCTGACGCGCTCGCGGCAATCGTGGAGGCGCAGGCG 2760
 DB 2989 GGTGACCAAGAGTCAAGAGATGACCTGACGCGCTCGCGGCAATCGTGGAGGCGCAGGCG 3048
 QY 2761 CCGCACCGCGGCAACAATGTGTGCACTGTGTGCGACCGGCAACAACGAGCTTCTCATG 2820
 DB 3049 CCGCACCGCGGCAACAATGTGTGCACTGTGTGCGACCGGCAACAACGAGCTTCTCATG 3108
 QY 2821 GTGGAATTGACATCGCAATGAGAAATCAGTTGTCTGGAACCTGCTGAGGAATTTCAAA 2880
 DB 3109 GTGGAATTGACATCGCAATGAGAAATCAGTTGTCTGGAACCTGCTGAGGAATTTCAAA 3168

QY 2881 AACAGCAAGGGTGGCAGAGAGCTGTGGTGTGTTTCAAACTTGTGCTGTCTTCTAC 2940
 DB 3169 AACAGCAAGGGTGGCAGAGAGCTGTGGTGTGTTTCAAACTTGTGCTGTCTTCTAC 3228
 QY 2941 AAATCACACGAGCAATCATCCCTTGGCAGCTGCTCTGTGCTGCTACTGCTACG 3000
 DB 3229 AAATCACACGAGCAATCATCCCTTGGCAGCTGCTCTGTGCTGCTACTGCTACG 3288
 QY 3001 ATCCCTCTGAGTCCGAGAACTCCAGAAAGCTACGTGTTCAAGTGAACATTCAAGTCC 3060
 DB 3289 ATCCCTCTGAGTCCGAGAACTCCAGAAAGCTACGTGTTCAAGTGAACATTCAAGTCC 3348
 QY 3061 CACGTCTACTCTTCAAGCGGAGAAAGCAGTACAGTTCGAAAGGATGGAAGGATC 3120
 DB 3349 CACGTCTACTCTTCAAGCGGAGAAAGCAGTACAGTTCGAAAGGATGGAAGGATC 3408
 QY 3121 CGCAGTCCACAGCTCTGCTCTGCGACCCCAAGTGTGAGCCACAAGAGTCTTGTG 3180
 DB 3409 CGCAGTCCACAGCTCTGCTCTGCGACCCCAAGTGTGAGTCAAAAGAGTCTTGTG 3468
 QY 3181 TATGATGAGCGGAGACACTGCTTTCCGAGTGGCTGCTTCTGGAAGACGTTCTCTT 3240
 DB 3469 TATGATGAGCGGAGACACTGCTTTCCGAGTGGCTGCTTCTGGAAGACGTTCTCTT 3528
 QY 3241 CTTCTGTATTAATGAAGCTTGGTAAATTAACACCTGTCTGAAAAATCAAAACATGCTT 3300
 DB 3529 CTTCTGTATTAATGAAGCTTGGTAAATTAACACCTGTCTGAAAAATCAAAACATGCTT 3588
 QY 3301 CCACAGAGCTCTCTGTCTCCACAGCGGCTTTTAAACCCGACCTCTCAAGCTTTGAA 3360
 DB 3589 CCACAGAGCTCTCTGTCTCCACAGCGGCTTTTAAACCCGACCTCTCAAGCTTTGAA 3648
 QY 3361 TGAACAGGCTCCCACTTCCATCTCGAGATCCGCTGGGGGGGCTGTTTAAAGCTAGTG 3420
 DB 3649 TGAACAGGCTCCCACTTCCATCTCGAGATCCGCTGGGGGGGCTGTTTAAAGCTAGTG 3708
 QY 3421 CCAGTATTAACATTTGCTATT 3442
 DB 3709 CCAGTATTAACATTTGCTATT 3730

RESULT 4
 ID AAA08582
 AC AAA08582; standard; DNA; 4687 BP.
 XX
 DT 19-JUL-2000 (first entry)
 XX
 DE Human cytoskeleton associated protein 2 (CYSKP-2) coding sequence.
 XX
 KW Cytoskeleton associated protein; CYSKP-2; cancer; proliferative;
 KW autoimmunity; inflammatory; vesicle trafficking; neurological;
 KW cardiovascular; cell motility; reproductive; muscle disorder; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 90..3227
 FT /tag= a
 FT /product= "CYSKP-2"
 XX
 PN MO200017355-A2.
 XX
 PD 30-MAR-2000.
 XX
 PF 17-SEP-1999; 99WO-US021565.
 XX
 PR 18-SEP-1998; 98US-0172226P.
 PR 27-APR-1999; 99US-0131321P.
 XX
 PA (INCY-) INCYTE PHARM INC.
 XX

Db	1722	GCTAAGAAAGTGTCTACACCGAGGAACTATCTGAAGATCTCGAAGTTATCACTTCG	1781
Qy	1741	TGTTTTCAAGCACTAGTACGAAAGAGACGCTCATGCTCCGGAAGCACTGAAAAGTCTCAT	1800
Db	1782	TGGTTTTAGAACACAGTAGCAAGAGAGACCGCACTGCGGAAGCACTGAAAAGTCTCAT	1841
Qy	1801	TTCCCGAATTTTGAACCTTTTGCACAAAATTTGATCTATTTTCTCAGAAATTTGAGCA	1860
Db	1842	TTCCCGAATTTTGAACCTTTTGCACAAAATTTGATCTATTTTCTCAGAAATTTGAGCA	1901
Qy	1861	CGACTTGCCTGTGGGAAGCGCGCTCAATGCTCCCAATCAGAATTCACAAAGATCGGC	1920
Db	1902	CGACTTGCCTGTGGGAAGCGCGCTCAATGCTCCCAATCAGAATTCACAAAGATCGGC	1961
Qy	1921	GATGTGATGTGAAGAACATTCAAGGCTAGAGCACTCTGCGGCTCACTGTGAAAGCAC	1980
Db	1962	GATGTGATGTGAAGAACATTCAAGGCTAGAGCACTCTGCGGCTCACTGTGAAAGCAC	2021
Qy	1981	AGCGAGGCTTTGGAGGCGCTTGAGAAATGGAAATCAAGAGCTCCGCGGCTTGAGAACTTC	2040
Db	2022	AGCGAGGCTTTGGAGGCGCTTGAGAAATGGAAATCAAGAGCTCCGCGGCTTGAGAACTTC	2081
Qy	2041	TGCAAGACTTTTGAAGCTCAGAAAGGTGTTATACCAACGCTCAACACTTCTCTGTGCG	2100
Db	2082	TGCAAGACTTTTGAAGCTCAGAAAGGTGTTATACCAACGCTCAACACTTCTCTGTGCG	2141
Qy	2101	CCACTGCACCGGCTCATGCACTACAGAGAGTCTTGAGCGGCTGTGCMAACACACCGG	2160
Db	2142	CCACTGCACCGGCTCATGCACTACAGAGAGTCTTGAGCGGCTGTGCMAACACACCGG	2201
Qy	2161	CCGAGCCACGCGCACTTCAAGGACTGCGGAGCGCTTTGGCAGATGACGAGAAATGGTG	2220
Db	2202	CCGAGCCACGCGCACTTCAAGGACTGCGGAGCGCTTTGGCAGATGACGAGAAATGGTG	2261
Qy	2221	GCACAGCTCCACGGTACGATGATCAAGATGAGAAATTTCCAGAAAGCTGCAGAACTCAAG	2280
Db	2262	GCACAGCTCCACGGTACGATGATCAAGATGAGAAATTTCCAGAAAGCTGCAGAACTCAAG	2321
Qy	2281	AAAGATTGATTTGGCATTTGACATCTTGTGGTTCCGGGAAGGAGTTCACTCGTCTGGGC	2340
Db	2322	AAAGATTGATTTGGCATTTGACATCTTGTGGTTCCGGGAAGGAGTTCACTCGTCTGGGC	2381
Qy	2341	AGCTCAGCAAGCTCTCGGGGAAGGGGCTCCAGACGCGATGTTCTTCTGTTCAAGCAC	2400
Db	2382	AGCTCAGCAAGCTCTCGGGGAAGGGGCTCCAGACGCGATGTTCTTCTGTTCAAGCAC	2441
Qy	2401	GTCCTGCATTTCAAGAGCGGGGGCTGACGGGCTCCAAATCAAGTTTAAAGTCCAGGGCAG	2460
Db	2442	GTCCTGCATTTCAAGAGCGGGGGCTGACGGGCTCCAAATCAAGTTTAAAGTCCAGGGCAG	2501
Qy	2461	CTCCCGCTCTATGGCATATGCAATTTGAGAGAGCAAGACGAGTGGGGGTTCCCACTGC	2520
Db	2502	CTCCCGCTCTATGGCATATGCAATTTGAGAGAGCAAGACGAGTGGGGGTTCCCACTGC	2561
Qy	2521	CTGACCTCTCGGGGCGACGGGCAGTCCATCTGTGGCGCGCAAGTCTTCTGGTCCGAGATG	2580
Db	2562	CTGACCTCTCGGGGCGACGGGCAGTCCATCTGTGGCGCGCAAGTCTTCTGGTCCGAGATG	2621
Qy	2581	GAGAAAGTGGGTTTGAAGACATCCAGATGGCCATTGACTGTGGGGGAGAAAGACAGCCCC	2640
Db	2622	GAGAAAGTGGGTTTGAAGACATCCAGATGGCCATTGACTGTGGGGGAGAAAGACAGCCCC	2681
Qy	2641	GCCTCTGAGTTCTTGGGCGAGAGCCCCCTTACAAACAAGTCCCTGATGAGACCAAGCGG	2700
Db	2682	GCCTCTGAGTTCTTGGGCGAGAGCCCCCTTACAAACAAGTCCCTGATGAGACCAAGCGG	2741
Qy	2701	GCTGACCAAGAGTCAAGAGATGACCTTGAGCGCTTGTGCGCACATCGCTTGAAGCGCACGCC	2760
Db	2742	GCTGACCAAGAGTCAAGAGATGACCTTGAGCGCTTGTGCGCACATCGCTTGAAGCGCACGCC	2801
Qy	2761	CCGCAACCGCGGCAACAAATGTGTGACGTGTGTGTGGCACCGCAACCAAGGTTCTTCATG	2820
Db	2802	CCGCAACCGCGGCAACAAATGTGTGACGTGTGTGTGGCACCGCAACCAAGGTTCTTCATG	2861

QY	2821	GTGACCTTCAGACATCGCAGTGTGGAAATCAATGTTCTGTGAAAACCTGCTGAGAAATTCAAA	2888
Db	2862	GTTGACTTCAGACATCGCAGTGTGGAAATCAATGTTCTGTGAAAACCTGCTGAGAAATTCAAA	2921
QY	2881	AACAGCAACGGGTGGCAGAACTGTGGTGTGTTCACAAACTCTGTGCTTTCTTCTAC	2940
Db	2922	AACAGCAACGGGTGGCAGAACTGTGGTGTGTTCACAAACTCTGTGCTTTCTTCTAC	2981
QY	2941	AAATCACACGAGCAATCATCCCTTGTCCAGCTGCTCTGTGTGGCTACTCGCTCAC	3000
Db	2982	AAATCACACGAGCAATCATCCCTTGTCCAGCTGCTCTGTGTGGCTACTCGCTCAC	3041
QY	3001	ATCCCTCTGAGTCGAGAAATATCCAGAAACATACGTTCAAGCTGCACTTCAAGTCC	3060
Db	3042	ATCCCTCTGAGTCGAGAAATATCCAGAAACATACGTTCAAGCTGCACTTCAAGTCC	3101
QY	3061	CACGCTCTACTCTCAGGGCGGAAAGCAGTACAGTTCCGAAAGGTGATGGAAGTATC	3120
Db	3102	CACGCTCTACTCTCAGGGCGGAAAGCAGTACAGTTCCGAAAGGTGATGGAAGTATC	3161
QY	3121	CGCAGTGCACCAAGCTCTGCTCGCGAACCCCACTGTTAGCCACAAAAGTCTCTGTG	3180
Db	3162	CGCAGTGCACCAAGCTCTGCTCGCGAACCCCACTGTTAGTCAAAAAGTCTCTGTG	3221
QY	3181	TATGATGTGGCGGACACATCGTTTCCGAGTGGCTGCTTCCGGAAGAGTTTCCCTT	3240
Db	3222	TATGATGTGGCGGACACATCGTTTCCGAGTGGCTGCTTCCGGAAGAGTTTCCCTT	3281
QY	3241	CTTCTGTATTATGAAAGCTGATAAAATTAAACACTGTCTGAAATCAAAAACATGAGCT	3300
Db	3282	CTTCTGTATTATGAAAGCTGATAAAATTAAACACTGTCTGAAATCAAAAACATGAGCT	3341
QY	3301	CCGACGAGCTTCTCTGTCTCCACAGCGCGCTTTTAAACCCGACCTCTCAGCGTTGAA	3360
Db	3342	CCGACGAGCTTCTCTGTCTCCACAGCGCGCTTTTAAACCCGACCTCTCAGCGTTGAA	3401
QY	3361	TGAAACAGGCGTCCACACCTCCAGCTCTGGCATCGCTGGGGGGCGCTGTTCTTAGCTAGT	3420
Db	3402	TGAAACAGGCGTCCACACCTCCAGCTCTGGCATCGCTGGGGGGCGCTGTTCTTAGCTAGT	3461
QY	3421	CCAGTATTAAAAACATTGTCATT	3442
Db	3462	CCAGTATTAAAAACATTGTCATT	3483
RESULT 5			
ID	AA564760	standard; cDNA; 3187 BP.	
XX	AA564760;		
AC	AA564760;		
DT	13-FEB-2002	(first entry)	
DE	DNA encoding novel human diagnostic protein #564.		
KM	Human; chromosome mapping; gene mapping; gene therapy; forensic;		
KN	food supplement; medical imaging; diagnostic; genetic disorder; ss.		
OS	Homo sapiens.		
PN	WO200175067-A2.		
PD	11-OCT-2001.		
PF	30-MAR-2001; 2001WO-US008631.		
PR	31-MAR-2000; 2000US-00540217.		
PR	23-AUG-2000; 2000US-00649167.		
PA	(HYSE-) HYSEQ INC.		
PI	Drmanac RT, Liu C, Tang YT;		

XX WPI: 2001-639362/73.
DR P-PSDB: ABG00573.
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 1; SEQ ID NO 564; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotide are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SO Sequence 3187 BP; 810 A; 874 C; 876 G; 627 T; 0 U; 0 Other;
Query Match 92.1%; Score 3170.2; DB 5; Length 3187;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3183; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
QY 1 CCCCCAGCGCGCGCGCGTGTGAGATATCTCTAAGCCGCTTCAATGAGAGAAATA 60
DB 1 CCCCCAGCGCGCGCGCGTGTGAGATATCTCTAAGCCGCTTCAATGAGAGAAATA 60
QY 61 GAGCAGAGCCGACCCACGATCACTGAGGCGCCCGGAAAATTGCGGAGTACGTACC 120
DB 61 GAGCAGAGCCGACCCACGATCACTGAGGCGCCCGGAAAATTGCGGAGTACGTACC 120
QY 121 TTGGAACGTGACAGAAAGCCGCCCAACACCTTCAGAAAACTCGTGTCAATAATC 180
DB 121 TTGGAACGTGACAGAAAGCCGCCCAACACCTTCAGAAAACTCGTGTCAATAATC 180
QY 181 CAGATGCTGAGTACACCCAGAGAGCATTTGAAGTTCCAAAGAGCTCTGGGAAAGTG 240
DB 181 CAGATGCTGAGTACACCCAGAGAGCATTTGAAGTTCCAAAGAGCTCTGGGAAAGTG 240
QY 241 CTGCTGATGATGATTTGACCACTCAACTCTGTGGAAGTGAATTTGGCCCTCAG 300
DB 241 CTGCTGATGATGATTTGACCACTCAACTCTGTGGAAGTGAATTTGGCCCTCAG 300
QY 301 TTTCTGATCAGAAAAAGATCAAGGTGTGCTGAGATCTCTAAACCCATTTGAAACAG 360
DB 301 TTTCTGATCAGAAAAAGATCAAGGTGTGCTGAGATCTCTAAACCCATTTGAAACAG 360
QY 361 ATTAGAAGCCAAAGACGTTGTGTTAAGTTGTGTAATTTCTTCCGCTGACAC 420
DB 361 ATTAGAAGCCAAAGACGTTGTGTTAAGTTGTGTAATTTCTTCCGCTGACAC 420
QY 421 ACACAACTCAAGAACTCAACAGTACTGTTCGCTGAGAGTGAAGAGACTTG 480
DB 421 ACACAACTCAAGAACTCAACAGTACTGTTCGCTGAGAGTGAAGAGACTTG 480
QY 481 GCTCAAGGAGGTGACGTGTATGACACGAGGACGCTCTTGAATTTCAACATTTG 540
DB 481 GCTCAAGGAGGTGACGTGTATGACACGAGGACGCTCTTGAATTTCAACATTTG 540

DB 481 GCTCAAGGAGGTGACGTGTATGACACGAGGACGCTCTTGAATTTCAACATTTG 540
QY 541 CAATCTGAGATTTGGGATTTTGTAGAGCTTTGACAGAGACACTTGACAAAATAA 600
DB 541 CAATCTGAGATTTGGGATTTTGTAGAGCTTTGACAGAGACACTTGACAAAATAA 600
QY 601 TACATACCTCAGACAGACGACATAGAGACAAATCGTGAATTTTCAACATACCAATT 660
DB 601 TACATACCTCAGACAGACGACATAGAGACAAATCGTGAATTTTCAACATACCAATT 660
QY 661 GACAAACACACGACGAAATCAGATTTCCAGCTCTTGAAGATTCCTTCAACTGGGCC 720
DB 661 GACAAACACACGACGAAATCAGATTTCCAGCTCTTGAAGATTCCTTCAACTGGGCC 720
QY 721 TATGAAATCCGGTTACACCGGCGCAAGGACAGGAAAGGACAGATCAATCGGCCGT 780
DB 721 TATGAAATCCGGTTACACCGGCGCAAGGACAGGAAAGGACAGATCAATCGGCCGT 780
QY 781 GCCAACACGAGGAATCTAGTGTTCAGGGTTTCACTAAGATCAATGCTTCAACTGGGCC 840
DB 781 GCCAACACGAGGAATCTAGTGTTCAGGGTTTCACTAAGATCAATGCTTCAACTGGGCC 840
QY 841 AAGTCCGGAAGCTGAGCTTCAAGAGAAAGCGCTTTTCATACAGCTCCGCGCAATGCC 900
DB 841 AAGTCCGGAAGCTGAGCTTCAAGAGAAAGCGCTTTTCATACAGCTCCGCGCAATGCC 900
QY 901 AATAGTGGATACAGGATACCTTGAATTCCTGATGGGACAGCGGGAATTTGCAAGTCC 960
DB 901 AATAGTGGATACAGGATACCTTGAATTCCTGATGGGACAGCGGGAATTTGCAAGTCC 960
QY 961 TTTCTGAAAATCTGTGTTGAACATCATGCTCTTTTGAAGCTTTTGAAGGCCAAACCA 1020
DB 961 TTTCTGAAAATCTGTGTTGAACATCATGCTCTTTTGAAGCTTTTGAAGGCCAAACCA 1020
QY 1021 AAGCCCAAGCCGCTCTCTTTAGCCGGGGTCATCATTTCCGTTAGTGTGAGCTCAG 1080
DB 1021 AAGCCCAAGCCGCTCTCTTTAGCCGGGGTCATCATTTCCGTTAGTGTGAGCTCAG 1080
QY 1081 AAGCAGTTCTGACATATGTTAAAGAGAGACATTAAGAGAGTCACTTTGAAGAAAG 1140
DB 1081 AAGCAGTTCTGACATATGTTAAAGAGAGACATTAAGAGAGTCACTTTGAAGAAAG 1140
QY 1141 CACAGCAAGATTCATTCATCCGAGCCTTGTTCACAGCCTTACAGAACTGAATTCGAA 1200
DB 1141 CACAGCAAGATTCATTCATCCGAGCCTTGTTCACAGCCTTACAGAACTGAATTCGAA 1200
QY 1201 GTGCTGAGACGATCTCAGACAGACACAGCCTTACATTTGAGAGAAGTGTCC 1260
DB 1201 GTGCTGAGACGATCTCAGACAGACACAGCCTTACATTTGAGAGAAGTGTCC 1260
QY 1260 AAGGGGCGAGAGCTGCGGCGAGAAAGAAAGGTTTCCGCGGGGAGCCGGGGTCC 1319
DB 1260 AAGGGGCGAGAGCTGCGGCGAGAAAGAAAGGTTTCCGCGGGGAGCCGGGGTCC 1319
QY 1320 GACCCGAGCCCTGCGCGAGAGAAAGCCCGCGGTTAAACAAGCGGAGCGAGCCG 1379
DB 1320 GACCCGAGCCCTGCGCGAGAGAAAGCCCGCGGTTAAACAAGCGGAGCGAGCCG 1379
QY 1380 GACCCGAGCCCTGCGCGAGAGAAAGCCCGCGGTTAAACAAGCGGAGCGAGCCG 1439
DB 1380 GACCCGAGCCCTGCGCGAGAGAAAGCCCGCGGTTAAACAAGCGGAGCGAGCCG 1439
QY 1440 TCAGCCCGCAGACGACACAGGCTCCCTGACTGGAGTCTTCACTTTCCGAGCTGTC 1499
DB 1440 TCAGCCCGCAGACGACACAGGCTCCCTGACTGGAGTCTTCACTTTCCGAGCTGTC 1499
QY 1500 TGTGAATCTGCAAGGAGAGTGTGCTTCCCAACGTTGCTCCCAACTGAGCCC 1559
DB 1500 TGTGAATCTGCAAGGAGAGTGTGCTTCCCAACGTTGCTCCCAACTGAGCCC 1559
QY 1560 CGACACCAAGAGGCTCTCTCTTGAATGACAGGCTGTGTAATGACAGGCTGTGCCC 1619
DB 1560 CGACACCAAGAGGCTCTCTCTTGAATGACAGGCTGTGTAATGACAGGCTGTGCCC 1619
QY 1620 CGACACCAAGAGGCTCTCTCTTGAATGACAGGCTGTGTAATGACAGGCTGTGCCC 1620
DB 1620 CGACACCAAGAGGCTCTCTCTTGAATGACAGGCTGTGTAATGACAGGCTGTGCCC 1620

QY 1620 GACGAGCATGAGGATGAGGCGCGGAGGAGAGATTCCCAACTGATTAAGCGTACTTCAAT 1679
DB 1621 GACGAGCATGAGGATGAGGCGCGGAGGAGAGATTCCCAACTGATTAAGCGTACTTCAAT 1680
QY 1680 AGCTAAGGAAGTGTCTACCAACGAGCGAATATCTGAAGATCTCGAAGTTATCACTTC 1739
DB 1681 AGCTAAGGAAGTGTCTACCAACGAGCGAATATCTGAAGATCTCGAAGTTATCACTTC 1740
QY 1740 GTGGTTTCAGAGCAGCTGAGCAAAAGAGAGCGCATGCGGAGACCTGAAAAGTCTCAT 1799
DB 1741 GTGGTTTCAGAGCAGCTGAGCAAAAGAGAGCGCATGCGGAGACCTGAAAAGTCTCAT 1800
QY 1800 ATTCCGGAATTTGAACCTTTGCAAAATTTCTACTAATTTCTGAAGAAATTTAGCA 1859
DB 1801 ATTCCGGAATTTGAACCTTTGCAAAATTTCTACTAATTTCTGAAGAAATTTAGCA 1860
QY 1860 AGCACTTGGCTGTGGGAGAGCGCTCAAAATGCCAAATCAGAGATTTACCAAGAAATCGG 1919
DB 1861 AGCACTTGGCTGTGGGAGAGCGCTCAAAATGCCAAATCAGAGATTTACCAAGAAATCGG 1920
QY 1920 CGATGTCACTGAGAGAACTTCAAGGCGATGAAAGCACTGGCGGCTCACTGTGGAAGCA 1979
DB 1921 CGATGTCACTGAGAGAACTTCAAGGCGATGAAAGCACTGGCGGCTCACTGTGGAAGCA 1980
QY 1980 CAGCGAGGCTTGGAGAGCGCTGAGAGATGGAATCAAGAGCTCCGCGCGCTGAGAGACTT 2039
DB 1981 CAGCGAGGCTTGGAGAGCGCTTGGAGAGATGGAATCAAGAGCTCCGCGCGCTGAGAGACTT 2040
QY 2040 CTGCAAGAGACTTTGAGCTGAGAGAGTGTGTTACTTACCGCTCAACACTTCTCTGCGG 2099
DB 2041 CTGCAAGAGACTTTGAGCTGAGAGAGTGTGTTACTTACCGCTCAACACTTCTCTGCGG 2100
QY 2100 GCCACTGCAACCGGCTCATGCACTAACAGAGGTCTCTGAGCGGCTGTGCAACCAACCC 2159
DB 2101 GCCACTGCAACCGGCTCATGCACTAACAGAGGTCTCTGAGCGGCTGTGCAACCAACCC 2160
QY 2160 GCCGAGCCAGCGCGAATTCAGAGGAGCTCCGAGCGGCTTTGGCAGAGATCAACGAGATGGT 2219
DB 2161 GCCGAGCCAGCGCGAATTCAGAGGAGCTCCGAGCGGCTTTGGCAGAGATCAACGAGATGGT 2220
QY 2220 GGCACAGCTCCAACGATGATCAAGATGAGAAATTTCCAGAACTGCAACAACTCAA 2279
DB 2221 GGCACAGCTCCAACGATGATCAAGATGAGAAATTTCCAGAACTGCAACAACTCAA 2280
QY 2280 GAAAGATTGATTGGCATTGACAAATCTTGTGTTCCGGGAAAGGAGTTCACTCGTGGG 2339
DB 2281 GAAAGATTGATTGGCATTGACAAATCTTGTGTTCCGGGAAAGGAGTTCACTCGTGGG 2340
QY 2340 CAGCCTCAGCAAGCTCTCGGGGAAAGGCGTCCAGAGCGCGATTTCTTCTGTTCAACGA 2399
DB 2341 CAGCCTCAGCAAGCTCTCGGGGAAAGGCGTCCAGAGCGCGATTTCTTCTGTTCAACGA 2400
QY 2400 CGTCTGTATATACAGAGCGGCGGCTGACGCGCTCAATCAAGTTTAAAGTCCACGGGCA 2459
DB 2401 CGTCTGTATATACAGAGCGGCGGCTGACGCGCTCAATCAAGTTTAAAGTCCACGGGCA 2460
QY 2460 GCTCCCGCTCTATGCGATGACGATTGAGAGAGCGAAGACGAGTGGGCGCTGCCACTG 2519
DB 2461 GCTCCCGCTCTATGCGATGACGATTGAGAGAGCGAAGACGAGTGGGCGCTGCCACTG 2520
QY 2520 CTTGACCTTCGCGGCGAGCGGAGTCCATCATCTGCGCGCAAGTTCTCGGCTCCAGAT 2580
DB 2521 CTTGACCTTCGCGGCGAGCGGAGTCCATCATCTGCGCGCAAGTTCTCGGCTCCAGAT 2580
QY 2580 GGAGAAGTGGTTGAGAGATCCAGATGGCCATTGACCTGCGGAGAGAGCAGAGCCG 2639
DB 2581 GGAGAAGTGGTTGAGAGATCCAGATGGCCATTGACCTGCGGAGAGAGCAGAGCCG 2640
QY 2640 CGCCCTGAGTTCTTGCGCAGAGCGCCCTGCAACAAAGTCCCTGATGAAAGCCAGCG 2699
DB 2641 CGCCCTGAGTTCTTGCGCAGAGCGCCCTGCAACAAAGTCCCTGATGAAAGCCAGCG 2700

QY 2700 GGCTGACAGAGATCAGAGATGACCTGAGCGGCTCGGCACATCGCTGAGAGCGCAGGC 2759
DB 2701 GGCTGACAGAGATCAGAGATGACCTGAGCGGCTCGGCACATCGCTGAGAGCGCAGGC 2760
QY 2760 CCGCAGCGCGGCAACAAATGGTGCAGTGTCTGGCAGCCGCAACAGCAGCTTCCAT 2819
DB 2761 CCGCAGCGCGGCAACAAATGGTGCAGTGTCTGGCAGCCGCAACAGCAGCTTCCAT 2820
QY 2820 GGTGACCTTCAAGATGCGAGTGAAGATCAAGTTGTCTGGAACCTCTGAGAAATTTCAA 2879
DB 2821 GGTGACCTTCAAGATGCGAGTGAAGATCAAGTTGTCTGGAACCTCTGAGAAATTTCAA 2880
QY 2880 AAACAGCAACGAGTGCAGAAAGCTGTGGGTGGTGTCAAAATTTCTGCTGTCTCTA 2939
DB 2881 AAACAGCAACGAGTGCAGAAAGCTGTGGGTGGTGTCAAAATTTCTGCTGTCTCTA 2940
QY 2940 CAAATCAGACAGAGCAATATCCCTTGCAGACCTGCTCTGCTGAGTACTGCTCAC 2999
DB 2941 CAAATCAGACAGAGCAATATCCCTTGCAGACCTGCTCTGCTGAGTACTGCTCAC 3000
QY 3000 CATCCCTCTGAGTCCGAGAACATCCAGAAAGCTACGTTCAAGCTGACATTAAGTC 3059
DB 3001 CATCCCTCTGAGTCCGAGAACATCCAGAAAGCTACGTTCAAGCTGACATTAAGTC 3060
QY 3060 CCACGCTCTACTCTTCAAGGCGGAGAGCGAGTACAGTTCAAGAGGTGATGATGAT 3119
DB 3061 CCACGCTCTACTCTTCAAGGCGGAGAGCGAGTACAGTTCAAGAGGTGATGATGAT 3120
QY 3120 CCGCAGTGCACCAAGCTCTGCTCGGAGCCCAAGTGTGAGGCAAAAGATCTTGT 3179
DB 3121 CCGCAGTGCACCAAGCTCTGCTCGGAGCCCAAGTGTGAGGCAAAAGATCTTGT 3180
QY 3180 GTATTGA 3186
DB 3181 GTATTGA 3187

RESULT 6
AAC98992
ID AAC98992 standard; cDNA, 3094 BP.
XX
AC AAC98992;
XX
DT 09-MAR-2001 (first entry)
XX
XX Human pancreatic cancer antigen nucleotide sequence SEQ ID NO:220.
DE
XX Human; pancreas; pancreatic cancer; pancreatic cancer antigen; detection;
KW diagnosis; identification; cytostatic; neuroprotective; nootropic;
KW immunomodulatory; relaxant; contraceptive; gynaecological;
KW antiinflammatory; cardiant; gene therapy; chromosome mapping;
KW linkage analysis; tissue identification; tissue typing; forensic; neural;
KW immune system; muscular; reproductive; gastrointestinal; pulmonary;
KW cardiovascular; renal; proliferative; ss.
XX
OS Homo sapiens.
XX
PN WO20005320-A1.
XX
PD 21-SEP-2000.
XX
PF 08-MAR-2000; 2000WO-US005989.
XX
PR 12-MAR-1999; 99US-0124270P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Ruben SM;
XX
XX WPI; 2000-579444/54.
XX
XX P-PSDB; AAB54227.
XX
XX New nucleic acid that is a pancreatic cancer antigen for preventing,

PT treating, or ameliorating a medical condition, particular pancreatic
PT cancer, or for use in assays for diagnosing a pathological condition.
XX
PS Claim 1, Page 664-665; 1379pp; English.

XX AAC98773 to AAC99231 encode the human pancreatic cancer associated
CC proteins, called pancreatic cancer antigens, given in AAB54008 to
CC AAB54466. The human pancreatic cancer antigens have cytostatic,
CC neuroprotective, neurotropic, immunomodulatory, relaxant, contraceptive,
CC gynecological, cardiant and antiinflammatory activities, and can be used
CC in gene therapy. The polynucleotide and proteins can be used for
CC preventing, treating, or ameliorating a medical condition or in assays
CC for diagnosing a pathological condition or a susceptibility to one in a
CC subject. Binding partners to the proteins and the activity of the
CC proteins can be identified. The pancreatic cancer antigens can be used to
CC detect, treat or prevent pancreatic disorders, especially cancer.
CC Agonists and antagonists to the antigens can be screened for. The
CC pancreatic cancer antigen polynucleotides can be used to design nucleic
CC acid hybridization probes that can be used in chromosome mapping, linkage
CC analysis, tissue identification and/or typing and a variety of forensic
CC and diagnostic methods. The proteins can be used to generate antibodies
CC which are used to purify, detect and target the polypeptides, including
CC both in vivo and in vitro diagnostic and therapeutic methods. The
CC proteins can be used to treat or prevent neural, immune system, muscular,
CC reproductive, gastrointestinal, pulmonary, cardiovascular, renal or
CC proliferative disorders. AAC99232 to AAC99240 and AAB54467 represent
CC sequences used in the exemplification of the present invention
XX

SQ Sequence 3094 BP; 736 A; 849 C; 755 G; 749 T; 0 U; 5 Other;

Query Match 49.1%; Score 1688.6; DB 3; Length 3094;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 1699; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

QY 1740 GTGGTTTCAGACACAGTACGAAAGAGACCGCATGCGGAAAGCATTAAGTCTCAT 1799
DB 62 GTGGTTTCAGACACAGTACGAAAGAGACCGCATGCGGAAAGCATTAAGTCTCAT 121
QY 1800 ATTCCGAAATTTGAACCTTTCACAAATTTCAATCTAATTTTCCAAAGAAATTGAGA 1859
DB 122 ATTCCGAAATTTGAACCTTTCACAAATTTCAATCTAATTTTCCAAAGAAATTGAGA 181
QY 1860 ACGACTTGGCCTGTGGGAAGCGCGTCAAAATGCCAAATCAGAGATTACCAAGATCGS 1919
DB 182 ACGACTTGGCCTGTGGGAAGCGCGTCAAAATGCCAAATCAGAGATTACCAAGATCGS 241
QY 1920 CGATGTCATGCTGAAGAAATTCAGGGCATGAAGCACTGGCGCTCACTGTGGAAGA 1979
DB 242 CGATGTCATGCTGAAGAAATTCAGGGCATGAAGCACTGGCGCTCACTGTGGAAGA 301
QY 1980 CAGCGAGGCTTGGAGAGCCCTGGAAGATGGAATCAAGAGCTCCCGCGGCTGAGAACTT 2039
DB 302 CAGCGAGGCTTGGAGAGCCCTGGAAGATGGAATCAAGAGCTCCCGCGGCTGAGAACTT 361
QY 2040 CTGCAGAGACTTTGAGCTGCAGAAAGGTGTGTTACCTACCGCTCAACACTTCTCTGCG 2099
DB 362 CTGCAGAGACTTTGAGCTGCAGAAAGGTGTGTTACCTACCGCTCAACACTTCTCTGCG 421
QY 2100 GGCATCTGACCGGCTCATGCACTAACAAGAGTCTGTGAGCGGCTGTGCAACACACCC 2159
DB 422 GGCATCTGACCGGCTCATGCACTAACAAGAGTCTGTGAGCGGCTGTGCAACACACCC 481
QY 2160 GCGAGGCGCAAGCGGCACTTCAAGGAGCTGCGGAGCGCTTGGCAGAGATCAAGGATGCT 2219
DB 482 GCGAGGCGCAAGCGGCACTTCAAGGAGCTGCGGAGCGCTTGGCAGAGATCAAGGATGCT 541
QY 2220 GGCACAGCTCAAGGTATGATGATCAAGATGAGAAATTTCCAGAAAGCTGCAAGAACTCA 2279
DB 542 GGCACAGCTCAAGGTATGATGATGATGAGAAATTTCCAGAAAGCTGCAAGAACTCA 601
QY 2280 GAAAGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2339
DB 602 GAAAGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 661

QY 2340 CAGCTCAGCAAGCTCTGCGGGGAAGGGCTCCAGCAGGCAATGTTCTTCTTCAACGA 2399
DB 662 CAGCTCAGCAAGCTCTGCGGGGAAGGGCTCCAGCAGGCAATGTTCTTCTTCAACGA 721
QY 2400 CGTCTGCTATACAGAGCCGGGGCTGACGGCTTCAATCATGTTAAAGTCCAGGGCA 2459
DB 722 CGTCTGCTATACAGAGCCGGGGCTGACGGCTTCAATCATGTTAAAGTCCAGGGCA 781
QY 2460 GCTCCCGCTATATGCAAGATGATGAGAGCAAGACAGATGGGGGTGCCCACTG 2519
DB 782 GCTCCCGCTATATGCAAGATGATGAGAGCAAGACAGATGGGGGTGCCCACTG 841
QY 2520 CTTGACCTTCGGGGCCAGCGGCAAGTCCATCATCTGACCGGCAATTCGGTCCGAGAT 2579
DB 842 CTTGACCTTCGGGGCCAGCGGCAAGTCCATCATCTGACCGGCAATTCGGTCCGAGAT 901
QY 2580 GAGAAATGGGTTGAGACATTCAGATGCGCAATTCAGCTGGCGAAGAGCAGAGCC 2639
DB 902 GAGAAATGGGTTGAGACATTCAGATGCGCAATTCAGCTGGCGAAGAGCAGAGCC 961
QY 2640 CGCCCTGATTTCTGGCCAGCAGCCCTTGAACAAGTCCCTGATGAAAGCCAGCC 2699
DB 962 CGCCCTGATTTCTGGCCAGCAGCCCTTGAACAAGTCCCTGATGAAAGCCAGCC 1021
QY 2700 GAGTACAGAGATGAGAGATGAGCTGAGCGCTCGGCAATTCGTCGAGCGCAGG 2759
DB 1022 GAGTACAGAGATGAGAGATGAGCTGAGCGCTCGGCAATTCGTCGAGCGCAGG 1080
QY 2760 CCGGACCGCGGCAACATATGTCACGTGTGTCGACCGCAACACAGCTTCCAT 2819
DB 1081 CCGGACCGCGGCAACATATGTCACGTGTGTCGACCGCAACACAGCTTCCAT 1140
QY 2820 GGTGACCTTCAGCATCGAGTGGAGAAATCAGTTGCTGGAACCTGAGGAAATTC 2879
DB 1141 GGTGACCTTCAGCATCGAGTGGAGAAATCAGTTGCTGGAACCTGAGGAAATTC 1200
QY 2880 AAACAGCAACGGGTGGCAGAAAGCTGTGGTGTTCACAACTTCTGCTGTTCTA 2939
DB 1201 AAACAGCAACGGGTGGCAGAAAGCTGTGGTGTTCACAACTTCTGCTGTTCTA 1260
QY 2940 CAATACACAGAGCAATCATCCCTTCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2999
DB 1261 CAATACACAGAGCAATCATCCCTTCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320
QY 3000 CATCCCTCTGAGTCCGGAATCATCAGAAAGCTGTTCAAGCTGCACTTCAAGTC 3059
DB 1321 CATCCCTCTGAGTCCGGAATCATCAGAAAGCTGTTCAAGCTGCACTTCAAGTC 1380
QY 3060 CCAGCTCTACTTCAAGGGCGAAAGCGAGTACAGTTCGAAAGGTGAGAGTAT 3119
DB 1381 CCAGCTCTACTTCAAGGGCGAAAGCGAGTACAGTTCGAAAGGTGAGAGTAT 1440
QY 3120 CCGCAGTCCACAGCTTCTGCTGCGACCCCAAGCTGTTGAGCCCAAGAGTCTTGT 3179
DB 1441 CCGCAGTCCACAGCTTCTGCTGCGACCCCAAGCTGTTGAGCCCAAGAGTCTTGT 1500
QY 3180 GTATTGATGGCGGACACACTGTTTCGAGTGGCTTTCGGAAGAGTTCCTT 3239
DB 1501 GTATTGATGGCGGACACACTGTTTCGAGTGGCTTTCGGAAGAGTTCCTT 1560
QY 3240 TCTTCTGATTAATGAGCTGTTAAATTAACCTGCTGAAATCAAAAACATGCT 3299
DB 1561 TCTTCTGATTAATGAGCTGTTAAATTAACCTGCTGAAATCAAAAACATGCT 1620
QY 3300 TCCACAGCTCTCTGCTGCTCAAGCGCGTTTAAACCCCGACCTCTCAGCGTTGA 3359
DB 1621 TCCACAGCTCTCTGCTGCTCAAGCGCGTTTAAACCCCGACCTCTCAGCGTTGA 1680
QY 3360 ATGAACAGGCTCCCACTCCAGTCTGAGCATCCGCTGGGGGCGGTCTTCTTAGCTAGT 3419
DB 1681 ATGAACAGGCTCCCACTCCAGTCTGAGCATCCGCTGGGGGCGGTCTTCTTAGCTAGT 1740

QY 3420 GCCAGTATTAAACATTGTCATT 3442
|||
Db 1741 GCCAGTATTAAACATTGTCATT 1763
|||
RESULT 7
ADQ84807
ID ADQ84807 standard; cDNA; 3997 BP.
AC ADQ84807;
XX
XX 07-OCT-2004 (first entry)
XX
XX Human tumour-associated antigenic target (TAT) cDNA sequence #1621.
XX
XX human; tumour-associated antigenic target; TAT; cytostatic; gene therapy;
XX cancer; cell proliferative disorder; gene; ss.
XX Homo sapiens.
XX
XX MO2004060270-A2.
XX
XX 22-JUL-2004.
XX
XX 15-OCT-2003; 2003WO-US029126.
XX
XX 18-OCT-2002; 2002US-0418988P.
XX
XX (GETH) GENENTECH INC.
XX (WUTD/) WU T D.
XX (ZHOU/) ZHOU Y.
XX
XX Wu TD, Zhou Y;
XX
XX WPI; 2004-534300/51.
XX
XX New nucleic acid molecule and encoded polypeptide, for diagnosing,
XX preventing or treating cell proliferative disorders such as cancer.
XX
XX Claim 1; SEQ ID NO 1621; 5504bp; English.
XX
XX The present invention describes an isolated tumour-associated antigenic
XX target (TAT) nucleic acid comprising: (a) any of 4632 nucleotide
XX sequences (see SEQ ID NO.1 to 4632); (b) the full-length coding region of
XX (a); (c) the complement of (a) or (b); (d) a sequence that has 80%
XX sequence identity to (a)-(c); or (e) a sequence that hybridises to (a) -
XX (c). Also described: (1) an expression vector comprising the above
XX nucleic acid; (2) a host cell comprising the above expression vector; (3)
XX a process for producing a polypeptide; (4) an isolated polypeptide
XX comprising: (a) an amino acid sequence encoded by any of the above
XX nucleotide sequences; (b) an amino acid sequence encoded by the full-
XX length coding region of the above nucleotide sequences; or (c) a sequence
XX having at least 80% identical to (a) or (b); (5) a chimeric polypeptide
XX comprising the above polypeptide fused to a heterologous polypeptide; (6)
XX an isolated antibody that binds to the above polypeptide; (7) a process
XX for producing the antibody; (8) an isolated oligopeptide that binds to
XX the above polypeptide; (9) a tumour-associated antigenic target (TAT)
XX binding organic molecule that binds to the above polypeptide; (10) a
XX composition of matter comprising the above (chimeric) polypeptide,
XX antibody, oligopeptide or TAT binding organic molecule, in combination
XX with a carrier; (11) an article of manufacture comprising a container and
XX the composition of matter contained within the container; (12) methods of
XX inhibiting the growth of a cell that expresses the above protein, where
XX the growth of the cell is at least in part dependent upon a growth
XX potentiating effect of the above protein; (13) a method of
XX therapeutically treating a mammal having a cancerous tumour comprising
XX cells that express the above protein; (14) a method of determining the
XX presence of a protein in a sample suspected of containing the protein
XX described above; (15) methods of diagnosing the presence of a tumour in a
XX mammal; (16) a method for treating or preventing a cell proliferative
XX disorder associated with increased expression or activity of the above
XX protein; and (17) a method of binding an antibody, oligopeptide or
XX organic molecule to a cell that expresses the protein described above.

CC The TAT sequences have cytostatic activities, and can be used in gene
CC therapy. The composition and methods are useful for diagnosing,
CC preventing or treating cancer. The composition is also used for preparing
CC a medicament for the therapeutic treatment or diagnostic detection of a
CC cell proliferative disorder or cancer. The present sequence represents a
CC human TAT cDNA sequence from the present invention.
XX
XX Sequence 3997 BP; 998 A; 1132 C; 1068 G; 799 T; 0 U; 0 Other;
SQ
Query Match 30.1%; Score 1037.4; DB 12; Length 3997;
Best Local Similarity 60.2%; Pred. No. 1e-270;
Matches 1864; Conservative 0; Mismatches 1186; Indels 45; Gaps 7;
QY 69 GCCAGCCCAAGATCAACGACTGGGGCCCCGGAATAATGGGGGATAGTACTGTGAAGC 128
Db 150 GCAGACTGCAGGGATGCGCTGGGGTCCAGACCCCTGGGGGTTAGACCTTTAGACC 209
QY 129 TGGACAGAAAGCCGCCCAACACCTTCAGGAAATCTGTCCATCAAAATCCAGATGCT 188
Db 210 TGGGAGACTCTCTTGGCCAGAAATGCAAGAGAACCTGCACCTCAGAGTAAGCTGCT 269
QY 189 GATGACACCCAGAGAGCAATTTGAAGTCCCAAGAGCTCTGGAGAGTGTCTGGA 248
Db 270 GCACAAACACCATGGAATATTTGACATGAGCCTTAATGCAATGCCAGATATTACTGAC 329
QY 249 TGCAGTTGCAACCACTCAACCTCGTGAAGGTGATATTTGGCTCGAGTTCTGA 308
Db 330 ACAAGTGTGAAGCGTTTAAACCTGTGAAGATGACTACTCTGGAGTGAAGTTTCAAAA 389
QY 309 TCACAAAAGATCAACGATGCTGTGATCTCTTAAACCATTTGTGAACAGATTGAAG 368
Db 390 TACTGATCTTCTGTGATTTGGCTTGAACCTAATGAACCATCATATTAGCAATAGGAAG 449
QY 369 GCCAAGACAGTTGTTGTTAAGTTGTGTGAATTTCTTCCGCTGACCAACACAACT 428
Db 450 GCCAAAGAAATGTGTCTTCCCTAGCTGTAATAATTTTTCACATCTGTGTGAGCT 509
QY 429 CCAGAGAAAGATCAACAGTACTGTTCGCGCTGACAGTGAAGAGAGACTTGCTAAG 488
Db 510 ACAAGAGAAATATACAAATTAATCTTTGCTTGAACCTTAAGAGAGACTGTGGAAG 569
QY 489 CAGTTGAGGTGAATATGACACACGCACTCTTGTATTTACACATTTGTGCAATCTGA 548
Db 570 GCGTTTGACTGTGTGATGACACACAGCGGCTTTCAGATCTCTGAGATCGGA 629
QY 549 GATTTGGGATTTTGTATGAAGCTTTGACAGAGACCTTAGCAAAAATTAATATACATAC 608
Db 630 AATAGGAGATTAAGATGAAGACCTGGAACGAGAGCACTCAAAAGTGAACGAGATTTGCC 689
QY 609 TCAGCAAGACGACTAGAGAGACAAATCGTGAATTTCAACATAACCATTTGAGCAAAAC 668
Db 690 TGGCCAGCACTGCTTGAAGAAATTAAGATTTCAACAGAACAGTGGGCCAGAC 749
QY 669 ACCAGCAGATCAAGATTTTCAGCTCTTAGAGATTTGCCCTGAGATGATGAAT 728
Db 750 ACTGTGTAGTGTGATTTCCAGATGCTGCAAAATGCTCAAAAGTTGGAATGACGGCAT 809
QY 729 CCGGTTGCAACCCGCCAAGAGAGAGGAGGACAAAGATCAATCTGGCCGTTGCCAACAC 788
Db 810 CAGATTTCAATGAGCTTCTGACAGGGAGAGAACCAAGATTTCAATGGCAGTTTCCACAT 869
QY 789 GGGAAATTAAGTTTCAAGGTTTCACTAAGATCAATGCTTCAACTGGGCCAAGGTGG 848
Db 870 GGGGTACTGTGTTCAGAGGCAACCAAAATCAACATTTCAATGCTTCAAGGTCCG 929
QY 849 GAAGTGAAGCTTCAAGAGAGAGAGGCTTTCTCATCAAGCTCCGCGCAGATGCCAATAGTGC 908
Db 930 TAACTTACCTTCAAGAGAGAGAGATTTCTTATCAAACTTATCCAGAGTTTCAAGAAC 969
QY 909 GTACCAAGATTAATCTTGAATCTTGTATGAGCCAGTGGGATTTCTGCAAGTCTTCTGAAA 968
Db 990 TTACCAAGGACATTAATTAATTTTGTGGTGTAGTGAATGATGAATGAACCTTCTGAAA 1049

QY 969 AATCTGTGTGAACATCATGCTTCTTTAGACTTTTGAAGCCCAACAAAGCCCA 1028
DB 1050 GATTTGTGTGAATACACACCTTTTATGACTTTTGAACAACTTAAGCCAAAGCCAA 1109
QY 1029 GCCCTGCTCTTTAGCCGGGGGTATCATTTTGGTTCACTGGTCCGACTCAGAGAGGT 1088
DB 1110 AGCCGCTCTTTCAGCCGGGGCTCTCTTCAGATTCAGTGAAGAACTCAGAAACACT 1169
QY 1089 TCTCGACTATGTAAAGAGAGACATTAAGAAAGTGCAGTTTGAAGAGACACAGAA 1148
DB 1170 AGTAGATTTATTTCAAGACAGTGAATGAAGAAATTCATATGAAGAAAGCACAGCAA 1229
QY 1149 GATTCATTTCTATCCGAGGCTTGCTTACAGCTTACAGAACTGAATTCGAAATGCTGA 1208
DB 1230 GACCCGACGCTC-----GTTGAGCTCTGACTGACAGACTTACAA 1271
QY 1209 GCAGTCTCAGCAGACAGCCTTATCTTTGAGAAAGTGCAGATCTTCAGGGGGCCA 1268
DB 1272 ACAGAGCATCTCATTTCCCGAAGGATTTGAGAACTCTGCTCTCCCATCTTCAGCCGATGC 1331
QY 1269 GAGCTGCCGGCGAGAAAGAACCGAAAGTTTCGCGGGGAGCCGGGGTCCGACCCGAG 1328
DB 1332 CTTTATCTGCTCTCTCTCCCTCCACTGTGTTCCCTTGCTGCTGACAGATTTTAAGACAG 1391
QY 1329 CCTGCGCCGAGAGAAAGCCCGCGGTAAACAAGCGGAGCGAGCCGCTCGCGCC 1388
DB 1392 CAGCAGCTCCCTCACAGATCCCAAGTTTCTTACGTCAAGATCCAGCTGACAGAGAGCG 1451
QY 1389 CACGAGAAAGAGAGAGAG---TGGTTAAGATAGAGCCACAGAGATTAACCTTAGCC 1445
DB 1452 CAGTGAAGAGTGGCTGAGAGGCCCGACACACCATGGGCCAGCCCTCGGGCCCCCGC 1511
QY 1446 CCGCAGCCAAAGACAGAGCTCCCTGACTGAGTGCCTCACTTTCCGAGCTGCTGTGA 1505
DB 1512 ACTCGAGCTGTGCTCAGGCTTTTCCAGAAAGATCTTCAGCTTCTCCCTCCAGCCGGA 1571
QY 1506 CTGCGAGGGGAGAGTGGCCCTGCGCAACGTGACTTGTCTCCCACTGAGCCCGACAC 1565
DB 1572 GAGCCCCCTGAGTCTGAGCCCTGCACTTTCAGGTGCTT-----TGGGCCAGCTGA 1622
QY 1566 CAGCAGGCTCTCTCCCTTATCAGCCCCGCTGCTGAATGACAGGCTGCCCCCGACGGA 1625
DB 1623 ACAGGGCTATCCCACTCTGAGCTCTGTCTCACTGATGCTGAGGAGCGGGATGGA 1682
QY 1626 CGATGAGATGAGGGCCGAGAGAAAGATTTCCCACTGATTAAGCCTATCTATAGCTTA 1685
DB 1683 CTGCGAG---GAGCCAGACACAGGCGGTGCTGAGAGAGAGCTTATCTATAGCTCA 1739
QY 1686 GGAAGTGTATCACCGAGCGAATATGTGAAGATCTGAAGTTATCATCTTCTGTGTT 1745
DB 1740 AAGAGATTTCTGCTACAGAACGAACATACCTCAAGATTTAGAAATTAATACGTGTGTT 1799
QY 1746 TCAAGACAAAGTGAAGAAAGAGAGAGCGCCATGCGGAGACATGAAAGTCTATATCC 1805
DB 1800 CCGCAGCGAGTGGTGAAGAGAGCGCATGCTGACTGTGATGACCTGCTCTTCTC 1859
QY 1806 GATTTTGAACCTTTGSCAAATTTTCTACTAATTTTCTCAAGAAATGAGCAAGACT 1865
DB 1860 CAACATCATCCCATTAATGATTTCAAGAGCTTCTGCGGAGGTGAAGACAGAGCT 1919
QY 1866 TSCCTGTGGGAAGCGGCTCAATGCCCCAATCAGAG---ATTACAAAGATCGGCA 1922
DB 1920 GGCACCTCTGGGAAGGGCCCTCCAAAGCCACACAAAGGACGTCATCAAGATCGGGA 1979
QY 1923 TGTCTATGCTGAAGAAATTTCAAGGCACTGAAGCACTTGCGGCTCACTGTGAGACAG 1982
DB 1980 CATCTGCTCAGGAATCGCCAGTTAAAGGATTTACAGCTATCTTCAAGACATGA 2039
QY 1983 CGAGGCTTGAAGCCCTGGAAGATGAATCAAGAGCTCCCGCGGCTGAGAACTTCTG 2042
DB 2040 CGAGGCTTGAAGAACTGGAAGAGCTTACAAAGCTGTGAAGATTTGAAGGAGGTGA 2099
QY 2043 CAGAGACTTTGAGCTGCAAGAGGTGTATCTTACCGCTCAACACTTCTCTGCGGCC 2102

DB 2100 CAGAGATTTGAGCTGCAAGAAAGTGTGCTACTTGTCTCAACAGTTCTGCTAAGCC 2159
QY 2103 ACTGACCGGCTCATGCACTACAGAGATGCTGTGAGGGGCTGTGCAACACACCGCC 2162
DB 2160 CATCAAGGGGTGCTGCACTACCGCTGCTGTGCGCCCTATCGGACATTTACAGCCC 2219
QY 2163 GAGCCACGCCGACTTCAGGAGCTGCCGAGCCGCTTTTGGCAGAGATCAGAGATGTGTC 2222
DB 2220 CGGGCACCATGACTACGCTGACTGACATGAGCCCTGAAGGCATCAGAGGTGACAC 2279
QY 2223 ACAGTCCACGGTACAGATGATCAAGATGAATTTCCAAAGCTGACAGAACTCAAGAA 2282
DB 2280 CACACTACAGCACTTCTCATCCGCTGAGAACTTGAAGAGCTTAAGGACTCAGAG 2339
QY 2283 AGATTTGATTTGGCATTTGCAATCTTGTGTTCCGGAAAGGAGTTCATCCGCTGGGCG 2342
DB 2340 GGAACCTGTGGCATAGAGAACCTATTTGCTCTGAGAGGATTCATCCGAGGGCTG 2399
QY 2343 CCTCAGCAAGCTCTCGGGGAAAGGGCTCCAGCAGCGCATGTTCTTCTGTTCAACAGCT 2402
DB 2400 CTTTCAAGGCTCACCAAGAAAGGGCTGACAGAGAGATGTTTTTCTGTTCTCAGATAT 2459
QY 2403 CCTGCTATACAGAGCCGGGGCTGACGCTCCATCATGTTTAAATTCACGGGAGCT 2462
DB 2460 GTTGTCTATACAAACAAAGGATTTGAGAGACAGCACTTCCGATCCGGGGCTCTCT 2519
QY 2463 CCGCTCTATGCGCATGAGATTTGAGAGAGGAAACAGAGTGGGGGTGCCCATGCT 2522
DB 2520 TCCCTCCAAAGCATGCTGTGTGAAGAAAGTATACAGATGTGTTTCTTCCACTGTTT 2579
QY 2523 GACCTTCGGGGCCAGCGGAGTCCATCATGTGTCGCGCACTTCTGCTCGAGATGA 2582
DB 2580 CACCATTAACGGGCTCAGAAACAAATCGTGTGCAACAGCACTCGGCTGAAGAA 2639
QY 2583 GAAGTGGTTGAAGACATTCAGATGSCATTTGACTGTGCGAGAAAGACAGACGCCGC 2642
DB 2640 GAAGTGAAGTGTGACCTGAATCTCGCATCCAGACACAAAGATGGCGGTGAACGCG 2699
QY 2643 CCTGATTCCTGGCGACAG-----GCCCCCTGACAAAGTCCCTGATGAAGCCAC 2696
DB 2700 CCTTCACTGCGAGGCCGACCTGTGTGCACTGTCTCCCGAATCTCCCAAGAGATATC 2759
QY 2697 CGCGGCTGACAGAGTGAAGAGATGACCTGAGCCCTCGCGCATGCTGTGAAGCCCA 2756
DB 2760 TCTG---GAGCAGAGTCAAGAAATGATGTGTGGGTGTGCGCAGCTCCCTGAGGGCA 2816
QY 2757 GGGCCCGCAGCGCGGCAACCAATGCTGCAAGTGTGCTGCGCACCGCAACACAGGCTC 2816
DB 2817 TGGCAGACCGGGCCAAACCAATGCAAGTGTGCTGTGTAACCGAAACACAGGCTGT 2876
QY 2817 CATGTGACTTCAGCATCGCAGTGAAGATCAGTTGTCTGAAACCTGCTAGGAATTT 2876
DB 2877 CAGGCGAATCACAAGTGCAGCTGTGAGAAACAGCTTTCAAGATATCTGTCAAGAAAT 2936
QY 2877 CAAAAACAGCAACGGGTGGCAGAAAGCTGTGGTGTGTTCAAACTTCTGCTGTCTT 2936
DB 2937 CAAAAACAGTATGTGCTGAGAAAGCTCTGGGTGTGTTTACCACTTCTGTTGTCTT 2996
QY 2937 CTACAAATCACAACAGGACATATCCCTTGGCCAGGCTGTGCTGTGCTATCTGCT 2996
DB 2997 CTACAAATCATACTGAGATGACTACCACTGCGCCGCTCCGCTGCTGGGCTAAGCT 3056
QY 2997 CACATCCCTCTGAGTCCGAGAAATCATCAGAAAGACTACGTTTCAAGCTCAGCTTCA 3056
DB 3057 GAGCATCCCGAGAGAGGCGGATGATACAAAGACTATGTTTTCAGAGCTCAGTTCAA 3116
QY 3057 GTCCAGGCTTATCTTACGTTCAAGGCGAAAGCGAGTACGTTCAAGGTTGATGAAT 3116
DB 3117 ATCCACGCTTACTTCTTCCGGGCTGAGAGCAAGTACATTTGAAGGTGATGAGGT 3176
QY 3117 GATCCGAGTGCACAGACTGTGCTCGCGACCCC 3151

Db 3177 GATCAGGGGGCCAGAGCTCAGCCGGAGGGGCC 3211

RESULT 8
ADR25675
ID ADR25675 standard; DNA; 3997 BP.
XX ADR25675:
XX 21-OCT-2004 (first entry)
XX Breast cancer prognosis marker #1536.
XX de; breast cancer; prognosis; gene expression; diagnosis.
XX Homo sapiens.
XX WO2004065545-A2.
XX 05-AUG-2004.
XX 15-JAN-2004; 2004WO-US001100.
XX 15-JAN-2003; 2003US-00342887.
XX (ROSE-) ROSETTA INPHARMATICS LLC.
XX (NECA-) NETHERLANDS CANCER INST.
XX Van't Veer LJ, He Y;
XX MPI; 2004-593473/57.
XX Classifying a breast cancer patient according to prognosis comprises
PT determining the similarity between the level of expression of each of
PT five genes in a cell sample taken from patient, to control levels.
XX Disclosure; SEQ ID NO 1536; 226pp; English.

The invention relates to a method of classifying a breast cancer patient
CC according to prognosis by determining the similarity between the level of
CC expression of each of five genes for which markers are listed in the
CC specification, in a cell sample taken from the breast cancer patient, to
CC control levels of expression for each respective five genes to obtain a
CC patient similarity value. The methods are useful for classifying a breast
CC cancer patient according to prognosis. Kits and computer program products
CC are useful for data analysis using the diagnostic, prognostic and
CC statistical methods of the invention. This sequence corresponds to a
CC marker used in the method of the invention.

Seq Sequence 3997 BP; 998 A; 1132 C; 1068 G; 799 T; 0 U; 0 Other;

Query Match 30.1%; Score 1037.4; DB 13; Length 3997;
Best Local Similarity 60.2%; Pred. No. 1e-270;
Matches 1864; Conservative 0; Mismatches 1186; Indels 45; Gaps 7;

QY 69 GCGGACCCAGGATCAGACTGGGGGCCCCGGAATTCGGGATCAGTACCTTGAAGC 128
DB 150 GCGAGCTGACGGAGTGGCTTGGGTGCGCCAGACCCCTGGGAGTTAGCACCCTTAGCC 209

QY 129 TGGACAGAAAGCCGCCCAACACCTTCAGGAAACTGTGTCATCAAAATCAGATGCT 188
DB 210 TGGGCAAGACTCTTGTCCCAAGATGCAAGAAAGCACTGCACCTCAGATTAAGCTGCT 269

QY 189 GAGTGAACCCAGAGAGGACTTGAAGTTCACAAAGAGCTCTGGGAAGGTGCTGTGA 248
DB 270 GGACAAACACCATGAAATATTTGACATTTGAGCTTAATGCGATGGCCAGGTATTACTGAC 329

QY 249 TGGAGTTTGAACACCTCAACTGTGGAAGGTAATTTTGGCTGTGAGTTTCTCTGA 308
DB 330 ACAAGTGTGAACGTTTAAACCTGTGAGATGACTTGTGGAGTGAAGTTTCAAAA 389

QY 309 TCACAAAAGAGTACGGTGTGGCTGTGATCTCTAAAACCATTTGTGAAACAGATTGAAG 368

Db 390 TACTGACTCTACTGGAATTTGGCTTGAACTATGAAGAACCATCATTTAGCAAAATACGAAG 449

QY 369 GCCAAAGACGTTGTTGTTAAAGTTTGTGGAAATCTTTCCGCTGACACACAACT 428
DB 450 GCCAAAGAAATGTGTGCTTCGCTACGTGTAAATTTTTCACCTGATCTGTGTAGCT 509

QY 429 CCAAGAAAGAACTCACAAGGAACTGTTCGGCGCTGAGTGAAGCAGAGCTTGAGTCAAG 488
DB 510 ACAAGAAAGAAATATACAAAGATCTTGTGTGCTTGCAACTTAAGAGAGACTGCTGAAAG 569

QY 489 CAGTTGAGGTATATAGACACGACGACCTCTTGAATTTCAACATTTGTGAATCTGA 548
DB 570 GCGTTGACCTGTGCTGACACCAAGCGGCGCTTCAAGTCCATCTCCTGACGTGGA 629

QY 549 GATTGGGATTTTGAATGAAGCTTGTGACAGAGACCTTAAGCAAAATTAATATACATAC 608
DB 630 AATAGAGATTAACGATGAAGAACGTGACCGAGGACCTCAAGTGAAGAGATTTTGGC 689

QY 609 TCAGCAAGACGACATGAGAGCAAAATCGTGAATTTCAACATTAACCATTTGACAAAC 668
DB 690 TGGCCAGCAGCACTGCTTGAAGAGTACTAGAAATTCATCAGAGCACTGGGCCAGAC 749

QY 669 ACCAGCAGATCAGATTTCCAGCTCCTAGAGATTGCCGCTGAGATGATGATGAT 728
DB 750 ACTGTGAGTGGGATTTCCAGGTGCTCGAAATGTCTCGAAAGTTGAATGTACGGCAT 809

QY 729 CCGGTTGACCCGGCCAAAGACAGGGAAGGACAGAAATCAATCTGCGCTGCCAACAC 788
DB 810 CAGATTTCAATGGCTTCTGACAGGAGGAAGAACCAAGATTCACCTGAGTTTCCACAT 869

QY 789 GGGAAATTCAGTGTTCAGGCTTTCATTAAGATCAATGCTTCAACTGGGCCAAGGTGG 848
DB 870 GGGTGTACTGTGTTTCAAGGACACCAAAATCAACCTTCACTGCTCAAGTCCG 929

QY 849 GAAGCTGACCTTCAAGAGAAAGGCTTCTCACTCAAGCTCGGCGCATGCCAATAGTGC 908
DB 930 TAACTTAAGCTTCAAGAGAAAGATTTCTTATCAAACTTCAAGAGTTTCATGAGACC 989

QY 909 GTACCAAGATTCCTTGAATTTCTGATGGCCAGTGGGATTTCTGAAATCTTTCTGAA 968
DB 990 TTACCAAGGACAACTTGAATTTTGTGGGTGATGAATGATGAATGAATCACTTCGAA 1049

QY 969 AATCTGTGTTGAACATCAAGACCTTCTTGAACCTTTTGAAGGCCCAACCAAGGCCAA 1028
DB 1050 GATTTGTGAGATTAACAACCTTTTGAACCTTTTGAACCACTTAAGCCAAAGCAAA 1109

QY 1029 GCCCGTCTCTTTAGCCGGGGGTCAATCTTCCGTGAGTGTGAGACTCAGAAAGCAGGT 1088
DB 1110 AGCCGTCTTCTCAGCCGGGGGTCTCTCTCAGATTAAGTGAAGAACTCAGAAACACT 1169

QY 1089 TCTGACTATGTTAAAGAGAGACATTAAGAAAGTGTGCACTTTGAAGAGACACGAA 1148
DB 1170 AGTAGATTTATTCAAAGACAGTGAATGAAGAAATTCATATGAAGAAAGCACAGCAA 1229

QY 1149 GATTCATTCATCCGAGGCTTGTCTCAGACCTACAGAACTGAATTCGAAAGTGTGGA 1208
DB 1230 GACCCACAGCTC-----GTTGAGCTGTGACTGCAAGCTTACCAA 1271

QY 1209 GCAGTCTCAGACAGACCAAGCCTTAATTTGGAAGAGGTGCCAATCTCAAGGGGCCA 1268
DB 1272 ACAGAGACATCTCATTCGCCGAGGATTTGAGACTCTCGCTCCCATCTTTCAGCGAATGC 1331

QY 1269 GAGCTCCGGCCGAGAAAGAAACCAAGGTTTTCGCGCGGAGACCGGGGTGCAACCGGAG 1328
DB 1332 CTTTAACTGTCTCTCCCTCACTCTGTGTCCCTCTGCGTGCAGAGATTTAAAGACAG 1391

QY 1329 CCTGCGCCGAGAGAAAGCCCGCGGGTAAACAAGACAGGAGCGGCTCTGCGGCGC 1388
DB 1392 CAGAGCTCTCTCAGAGATCTCCAGGTTTCTTAAGTCAAGAGATCCAGCTGCAAGAGGGG 1451

QY 1389 CACGAGAGAAAGAGAGAGG---TGTTAAGATAGAACCAACAGAGATTAACCTTCAGCC 1445
DB 1452 CAGTGAAGAGATGTGTGAGGCCCCGACACCAATGAGGCCACCTCGGGCCCCCGCC 1511

```

QY 1446 CCCGAGCCAAAGCAGAGCTCCCTGACCTGAGCTCTTCCAGACTGTGTGAA 1505
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1512 ACTCCAGCTTGATCCAGAGCTTTCCAGAAAGATCTTCAAGCTTCTCCCTCCAGCCGAA 1571
QY 1506 CTGCGAGGGGGAGTGGCCCTTGCCAACTGACCTTGTCTCCCACTGAGCCCGACAC 1565
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1572 GAGCCCCGTGAGCTTGAGCCCTGACATTTGAGGTGCTT-----TGGGCCAGCTGA 1622
QY 1566 CAGAGAGGCTCTCCCTTGTATGAGCCGCTGCTGAATGACAGAGCTGCCCCCGACGA 1625
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1623 AAGAGGCTATCCCACTCTGAGCCCTGTCTCACTGATGCTGAGCGAGCGGAGTGA 1682
QY 1626 CGATGAGATGAGGGCCGAGGAGAGATTTCCCACTGAATAAAGCTACTTATGACTAA 1685
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1683 CTGCGAG---GAGCCAGACACAGAGGGCGTGCCTGAGACGAGGCTTACTATAGTCA 1739
QY 1686 GGAAGTGTCTACCAAGAGCGAAATATTTGAAAGATCTGAAATTTATCATCTTGTT 1745
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1740 AAGAGATTCGTGCTACAGAAAGAACTCAAGATTTAGAAATTTATACCGTGTGTT 1799
QY 1746 TCAGAGCAGAGTGAAGAGAGAGAGCGCATGCGGAGCACTGAAAGTCTATATTTCC 1805
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1800 CCGCAGCGAGTGTGAAGAGAGAGCGCATGCTGCACTGTGATGACGCTGCTCTTC 1859
QY 1806 GAATTTGAACCTTGGCAAAATTTCTAATTTTCTCAAGAAATTTGAGCAAGACT 1865
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1860 CAACATCATCCCATTTATGATTTCCACAGAGGCTTCTGCGGAGGTGAGACAGAGCT 1919
QY 1866 TGCCTGTGGGAAAGCGGCTCAATGCCCAATCAGAG--ATTACAAAGATGGCGA 1922
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1920 GGCACCTGTGGGAAAGGGCCCTCCAAAGCCACAAAGGAGTCAATCAAGATCGGGGA 1979
QY 1923 TGTCACTGAGAAATTCATCAAGGAGTGAAGACCTGGGGCTCACCTGTGAGACAG 1982
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1980 CATCTGCTCAGAGAACATGCGCCAGTTAAAGATTTACAGCTTCTTCAAGACATGA 2039
QY 1983 CGAGGCTTTGAGAGCCCTGAGATGGAATCAAGAGCTCCGCGGCTGAGAACTTCTG 2042
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2040 CGAGGCTTCAACAGAACTGGAAAGGCTACCAAGCTGTAAAGAGTTGAGAGCAGTGA 2099
QY 2043 CAGAGACTTTGAGCTGAGAGAGGTGTGTTACCTACCGCTCAACACTTCTCTGCGGC 2102
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2100 CAAAGAGTTTGAAGCTGCAAGAGGTGTGTTACTTCAACACTTCTCTGTAAGCC 2159
QY 2103 ACTGCAACGGGCTCATGCACTAACAGAGTCTTGAGGCGGCTGTGCAACACACCGCC 2162
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2160 CATCCAGCGGCTGTGCACTACCGCTGTGCTGCGCCGCTATGCGACATTTACAGCCC 2219
QY 2163 GAGCCAGCGCGACTTCAGGGAATGCCAGCGGCTTTGAGAGATCACGAGATGTGAC 2222
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2220 CGGGGACCATGACTGATGCTGATGCCATGACGCCCTGAAAGCATACAGAGGTGACAC 2279
QY 2223 ACAAGCTCCAGGTAAGATGATCAAGATGAGAAATTTCCAGAACTGACGAACCTAAGAA 2282
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2280 CACACTTACGCACTTCTCATCGGCTGAGAACTGCAAGAACTTAACGAGCTGAGAGG 2339
QY 2283 AGATTGATTTGAGATTTGACAACTTTGTGTTCCGGGAGAGGAGTTTATCGTCTGGGAG 2342
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2340 GGAAGCTGTGGGATGAGAACTTCTCTCTGCGAGGAGTTTATCCGTGAGGGCTG 2399
QY 2343 CCTCAGCAAGCTCTCGGGGAGAGGGGCTCCAGCAGCGCATGTTCTTCTGTTCAAGAGCT 2402
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2400 CTTTCAAGAGCTCACAAAGAGGCGCTGCAGAGAGAGATGTTTTTCTGTTCAGATAT 2459
QY 2403 CCTGCTATACAGAGCCGGGGGCTGACGGCTTCGAATCGTTTAAAGTCCAGGGGAGCT 2462
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2460 GTTGTGTTGACAAAGAGAGGAGTTGCAAGGACCAAGCTTCCGAGATCCGGGGCTCTCT 2519
QY 2463 CCGGCTCTATGGATGAGATTTGAGAGAGAGCAAGTGGGGGAGTCCCACTGCT 2522
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2520 TCCCTCTCAAGGATGCTGTGTGAGAAAGATGATTAAGAGTGTCTGTTCCACTGTTT 2579

```

```

QY 2523 GACCTCCGGGGCCAGCGGAGTCCATCATGCTGAGCCGCAAGTTCTGATCCGAGATGA 2582
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2580 CACCATTTACGGGCTCAGAAAACAAATGCTGTGGCAGCGCACTCCGGCTGAGAAAGA 2639
QY 2583 GAAATGGGTTGAGGACATCCAGATGAGCCATTGACCTGGCGGAGAAAGACAGAGCCCGC 2642
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2640 GAAATGGATGTGTGACCTGAACTCCGATCCCAAGCAAGAGTGGCGGTGAGACAGGC 2699
QY 2643 CCTGAGTTCTGAGCAGCA-----GCCCCCTGACAAACAAAGTCCGATGAAAGCCAC 2696
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2700 CCTTCATCTGCCAGGCCGCACTGTGTGACTGTGTCCCCCAAGATCCCCCAAGAGTATC 2759
QY 2697 CGCGGCTACAGAGAGTGAAGATGATGAGCGCTTCCGCGCATGCTGCTGAGAGCCCA 2756
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2760 TCTG---GAGCAGAGAGTCAAGAAATGATGCTCGGGGTGTCCGCAAGCTCCCTGAGAGGGCA 2816
QY 2757 GGCCCCGACCGCGGCAACAAATGATGATGTGTGCTGAGCAGCGCAACCAAGGTCTC 2816
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2817 TGGCCAGACCGGGCCAAACCAATGACGTGTGTGTAACCGGAACACAGCGTGTG 2876
QY 2817 CATGTGACTTCAAGATCTGAGATGAGAAATCAATGATCTGGAACCTGTGAGGAAAT 2876
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2877 CAGGCGAATCAGAGTGAAGTGTGAGAACCAAGCTTTCAAGATATCTGTAAAGAAAT 2936
QY 2877 CAAAAACAGCAAGGGTGGCAGAAAGCTGTGGGTGTTCACAAATCTTGTGCTGTCTT 2936
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2937 CAAAAACAGTATGTGCTGGCAGAAAGCTGTGGGTGTTCACAAATCTTGTGCTGTCTT 2996
QY 2937 CTACAAATCAGACAGGACATCATCCCTTGGCAGGCTGTGCTGTGCTGTGCTGTGCT 2996
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2997 CTACAAACTCATTCAGATGACTACCACTGGCCAGGCTCCCGGTGCTGAGCTACAGCT 3056
QY 2997 CACATCCCTCTGTGAGTCCGAGAACATCCAGAAAGTCAAGTGTTCAGAGTCACTTCA 3056
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3057 GAGCATCCCGCAGGAGGCGGATGGATACAAAGACATATGTTTCAAGCTTCAAGTTCA 3116
QY 3057 GTCCACGTTCTACTACTTTCAGGGGCGAAAGCGAGTACAGTTTCAAGAGGTGATGAAT 3116
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3117 ATCCACGTTCTACTTCTTCCGGGCTGAGAGCAAGTACATTTGAAAGGTGATGAAGT 3176
QY 3117 GATCCGAGTGTGCACAGACTGTGCTCGGACCCC 3151
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3177 GATCCAGGGGCGCAGACAGCTCAGCCGAGAGGCC 3211
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 9
AD564452
ID AD564452 standard; DNA; 3997 BP.
XX
AC AD564452;
XX
AC 16-DEC-2004 (first entry)
XX
DT 16-DEC-2004 (first entry)
XX
DE Human KIAA0793 gene.
XX
KW Human; interferon; IFN; gene; ds; KIAA0793.
XX
OS Homo sapiens.
XX
PN US2004185489-A1.
XX
PD 23-SEP-2004.
XX
PF 17-MAR-2004; 2004US-00802432.
XX
PR 21-MAR-2003; 2003EP-00006263.
XX
PA (CERT/) CERTA U.
PA (FOSE/) FOSER S.
PA (WEYE/) WEYER K.
XX
PI CertA U, Foser S, Weyer K;
XX

```

DR MPI; 2004-689187/67.
 XX GENBANK; AB018336.
 PT Determination of biological activity of compound which can modulate gene
 PT transcription, involves contacting host with compound, determining
 PT general transcriptional gene response of host, and quantitating gene
 PT response induced by compound.
 XX
 PS Claim 8; SEQ ID NO 16; 53bp; English.
 CC The invention relates to a method for determining the biological activity
 CC of a compound which can modulate gene transcription. The method involves
 CC contacting a host with a compound, determining the general
 CC transcriptional gene response of the host and quantitating the gene
 CC response induced by the compound. This gene transcription assay method is
 CC useful for the determination of the biological activity of a compound
 CC which can modulate gene transcription. The present sequence is the human
 CC KIAA0793 gene. This sequence is used to estimate the influence of the
 CC peptidation site on the transcriptional activity of interferon (IFN) by
 CC measuring this IFN-induced gene expression pattern of the melanoma cell
 CC line ME35 via the oligonucleotide array technology.
 XX
 SQ Sequence 3997 BP; 998 A; 1132 C; 1068 G; 799 T; 0 U; 0 Other;
 Query Match 30.1%; Score 1037.4; DB 13; Length 3997;
 Best Local Similarity 60.2%; Pred. No. 1e-270;
 Matches 1864; Conservative 0; Mismatches 1186; Indels 45; Gaps 7;
 Oy 69 GCCGACCCGAGATCAAGCTGGGGGCCCGGAAATTCGGGGATCAATCTTGAACG 128
 Db 150 GCAGACTGAGGGATCGCTTGGGTGCCAGACCCCTGTGGAGATTAGCACTTGAAGCC 209
 Oy 129 TGGACGAAGACCGCCCGCCACACCTTCAGAAAATCTGTCTCATCAAAATCAAGTCT 188
 Db 210 TGGGCGACTCTCTTCCCGAGATGCAAGAGACACCTGCACCTAGATTAAGCTGCT 269
 Oy 189 GGATGACCCGAGAGGATTTGAATTCACAAAGATCTCTGGAGAGGTGCTGTGA 248
 Db 270 GGAACAACCATGGAATATTTCATTTAGCTTAATGGATGGCAGGATTAATCTGAC 329
 Oy 249 TGCAGTTTCACACCTCACTCGTGAAGGTGACTATTTTGGCTCGAGTTTCTGA 308
 Db 330 ACAGGTGGAAGCGTTTAACTGTGATGATGCTTAATGGATGGCAGGATTAATCTGAC 389
 Oy 309 TCACAAAAGATACGCTGTGCTGATCTCTAAACCATTTGTAACAGATTGAAG 368
 Db 390 TACTGCTCTATGATTTGGCTTGAACCTATGAACCATATTAAGGAATTAAGAA 449
 Oy 369 GCCAAGCAAGTGTGTTAAAGTTTGTGTAATCTTTCCGCTGACACACACACT 428
 Db 450 GCCAAGAAATGTGTCTTCTGCTTGAATTTTTCACCTGATCTGTGTGCTGCT 509
 Oy 429 CCAGAAGAACTCAAGGTAAGTGTGCTGCTGAGGTGAAGCAGAGCTTGTCTAAG 488
 Db 510 ACAGAAAGATATACAAATTAATTTGTTGCTTGAACCTTAAGAGACCTGTGGAAG 569
 Oy 489 CAGGTGAGCTGTAATGACACCGCAGCTCTTGAATTTCAACATTTGTAATCTGA 548
 Db 570 GCGTTTGAAGCTGTGCTGACACCAAGCGGCTTCTCAAGTCCATCTCTGAGTGG 629
 Oy 549 GATTGGGATTTTGAAGGCTTGAAGAGAGCACTTAGCAAAAATTAATATACATCC 608
 Db 630 AATTGAGATTTAGATGAAACGCTGACGAGAGCACTCAAGTGAAGATATTGCTCC 689
 Oy 609 TCAGCAAGACGACTAGAGCAAAATCGTGAATTTCAACATTAACACATTTGGAACAA 668
 Db 690 TGGCCAGCAGCACTGCTTGAAGAAATCAATCAATCAAGACGATGGGCGAGAC 749
 Oy 669 ACCAGAGATCAATGATTTCCAGCTCTTAAGATTTGCCGCTGAGATGATGTAAT 728
 Db 750 ACTGTGTGATCGGATTTCCAGGTGCTCGAAATTTGCTGAATTTGAAATGTAACGCAAT 809
 Oy 729 CCGGTTGACCCGCGCAAGGACAGGAAGCAAGATCAATCTGCGGTTGCCAAC 788

Db 810 CAGATTTCATATGCTTTCAGAGGAGAGAACCAAGATTCACTAGCTTTCCACAT 869
 Oy 789 GGGATTTAGTGTTCAGGGTTTCACTAAGATCAATGCTTCACTGGCCAGAGTGG 848
 Db 870 GGGTGTATCTGTGTTCCAGGGACCAACAAATCAACATTTCACTGTGTCCAGGTCCG 929
 Oy 849 GAAGCTAGCTTCAAGAGAGGCTTCTCATCAAGCTCCGGCCGATGGCCAAATAGTC 908
 Db 930 TAACTAAGCTTCAAGAGAAAGATTCTTATCAAACTTCACTCAAGGTTCAATGACC 989
 Oy 909 GTACCAAGATTCCTGAAATTCCTGATGCGAGTCCGGATTTCTGCAATCTTCTGAA 968
 Db 990 TTAACAGACATTAATTAATTTTGTGGGTGTAAGATGATGAATGAAGAACTTCTGAA 1049
 Oy 969 AATCTGTGTAACATCATGCTTCTTGAATCTTTTGAAGAGCCCAACCAAGCCCA 1028
 Db 1050 GATTGTGTGAGTATCAACACCTTTTGTAGACTTTTGAACCAACCTTAAGCCAAACAA 1109
 Oy 1029 GCGGCTCTTTTAAAGCGGGGGTCACTTTGCTGAGTGGGCTCAGAGAGGT 1088
 Db 1110 AGCGTCTTCTCAGCGGGGCTCTCTTCAAGATCAATGGAAGATCTAGAAACACT 1169
 Oy 1089 TCTGACTATGTTAAAGAGAGACATTAAGAGGTGCAATTGAAAGAGACACAGAA 1148
 Db 1170 AGTAAATTTTCAAGAGACAGTGAATGAAGAAATTCATATGAAGAGACAGCAA 1229
 Oy 1149 GATTCACTTATTCGAGGCTTGTTCACAGCTTACAGAACTGAATTCGAAGTCTGA 1208
 Db 1230 GACCCACAGTCC-----GTTCAAGCTTGAATGACAGACCTTACAA 1271
 Oy 1209 GGAATCTAGAGAGACAGACCTTATTTGGAAGAGGTGCGAATCTCCAGGGGCCA 1268
 Db 1272 ACAGAGCATCTAATTCCTCGAAGGATTGAAGACTCTGCTCCCATCTTCAAGCAATGC 1331
 Oy 1269 GAGCTCCCGGAGAGAAAGAAACCGAAGTTTCCGCGGGAGCGGGGTCCGACCCGAG 1328
 Db 1332 CTTTAACTGCTCTCTCCCTCACTCTGTGCTCCCTCTGCGCAGAGTTTAAAGACAG 1391
 Oy 1329 CCCTGCGCGAGAGAGAGCCCGGGGTAAACAGACAGCGAGAGAGCCCTCGGCGCC 1388
 Db 1392 CAGCAGCTCTTCAAGATCCCAAGGTTTCTAAGTCAAGATCCAGCTGACAGAGAGCG 1451
 Oy 1389 CACGAG 1445
 Db 1452 CAGTGAAGAGTGTGTGAG 1511
 Oy 1446 CCGGACCAAGACAGAGCTCCCTGACTGAGCTCTCACTTCCGAGCTGTCTGTGTA 1505
 Db 1512 ACTCAAGCTGTGTCAGAGCTTTCAGAGAGAGTCTTCAAGCTTCTCCCTCCAGCGGAA 1571
 Oy 1506 CTGCGAG 1565
 Db 1572 GAGCCCTGAGTGTGAG 1622
 Oy 1566 CAAGCAGAGCTCTCCCTTGAATGAGCCGCTGTGTAATGACAGGCTGCCCCGAGAGGA 1625
 Db 1623 AAGAGGCTATCCCACTCTAGAGCTGTCTCTCAAGTGTGAGAGAGAGAGAGAGAGAG 1682
 Oy 1626 CGATGAGATGAGAGAGAGAGAGAGAGATTTCCCACTGATTAAGGCTTCTATAGCTAA 1685
 Db 1683 CTGCGAG---GAGCCCAAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1739
 Oy 1686 GGAAGTGTCTACACCGAGAGCAATATCTGAAGATCTCGAAGTTATCACTTCTGTGTT 1745
 Db 1740 AAGAGATCTCGCTTACAG 1799
 Oy 1746 TCAAGACAG 1805
 Db 1800 CCGCAGCGAG 1859
 Oy 1806 GAATTTGAACCTTTCAGCAAAATTTCAATATTTTCTCAAGAGAAATTTGAGCAAGACT 1865

Db 1860 CAACATCCATCCCATCTATGATGATCCACAGAGGCTTCTCGCCGAGGTGAGCAGAGGCT 1919
Qy 1866 TGGCCCTGGGGAAGGCGGCTCAAAATGCGCAATTCAGAG--ATTACCAAGAAATGGGGA 1922
Db 1920 GGACCTCTGGGAAGGGGCTCCAAAGCCACACAAAGGAGCTCATCAACGATGGGGA 1979
Qy 1923 TGTATGCTGAAGAACATTCAGGGGCAATGAGCACTGGCGGCTCACTTGTGAGACAG 1982
Db 1980 CATCTGCTCAGGAACATGCGGCAAGTTAAAGAGTTTACAGCTACTTCCAAAGACATGA 2039
Qy 1983 CGAGGCTTTGAGGCGCTGGAGAAATGAATCAAGAGCTCCCGCGGCTGGAGAACTTCTG 2042
Db 2040 CAGGCTCTTAAGAACTGGAAGAGCTACAAACGCTTAAGAAAGTTGAGAGCAGGTGA 2099
Qy 2043 CAGAGCTTTGAGCTGAGAAAGTGTGTTACCAACGCTCAACACCTTCTCTGCGGCT 2102
Db 2100 CAAGAGTTTGAAGCTGCAAAAGGTCTGTACTTGTCTCAACAGTTCTCTGTGAAGCT 2159
Qy 2103 ACTGACCGGCTCATGACATCAAGCAGGTCCTGAGCGGCTGTGCAACACACCGGCT 2162
Db 2160 CATCCAGCGGCTGTGCTCACTACCGCTGTGCTGCGCGGCTATGCGACATTAAGAGCC 2219
Qy 2163 GAGCCACGCGCATCTTCAAGGACTGCGGAGCCGCTTTGGCAGAGATCAAGAGATGCTG 2222
Db 2220 CGGGGACCATGACTGCTGACTGCTGATGACGCTGAAAGCATCAAGAGGTGACAC 2279
Qy 2223 ACAGCTCCAGGTAAGATGATCAAGATGAGAAATTTCCAGAAAGCTGCAAGAACTCAAGAA 2282
Db 2280 CACACTACAGCAATTTCTATCCGGCTGAGAACTGCAAGAAAGCTTACCGAGCTGACAGG 2339
Qy 2283 AGATTGATTTGACATTTGATGATCTTGTGTTCCGGGAGGAGTTTCACTCTGCGGAG 2342
Db 2340 GGAACCTGTTGGGATTAAGAACTTCTGCTCTGCGGAGGAGTTTCACTCTGAGGAGCTG 2399
Qy 2343 CCTCAGCAAGCTTCTGCGGGAAGGGGCTCAAGCAGCGCATGTTCTTCTTCAAGAGCT 2402
Db 2400 CCTTCAAGAGCTCACCAAGAGGCGCTGCAAGAGAGATGTTTCTTCTTCAAGATAT 2459
Qy 2403 CCTGCTATCAGAGCGGGGCTGACGCGCTCCATCATGTTTAAAGTCCACGCGGAGCT 2462
Db 2460 GTTGTCTTACAGCAAGAAAGAGTTGCAAGGACAGCCACTTCCGAGTCCGGGCGCTCT 2519
Qy 2463 CCGGCTATGAGCATGACATTTGAGAGAGCAGAAAGAGTGGGGGTGCCCATGCT 2522
Db 2520 TCCGCTCCAAAGGCACTGTGTGAAGAAAGTGAATGAGTGTGTTCCACAGTGT 2579
Qy 2523 GACCTCCGCGGCGCAGCGGCACTCATCTGTGCGCGCAATTTCTCGGTCCGAGATGA 2582
Db 2580 CACCATCTACCGGCTCAGAAACCAATCTGTGTGCGAGCAGACACTCGCTGAGAAAGA 2639
Qy 2583 GAAAGTGGTGAAGCATCAGATGCGCATTTGACCTGGCGGAGAAAGAGAGAGAGCGCCGC 2642
Db 2640 GAAAGTGAATCTGAGACTGAATCTCGCGCATTCAGAGAGCAGAAAGAGTGGGAGTGAACG 2699
Qy 2643 CCCTGAGTCTTGCCAGCA-----GCCCGCTGCAACAAAGTCCCTGATGAAGCCAC 2696
Db 2700 CCCTGAGCTGCAAGGCGGCACTGTGTGCACTGTGCCCGCAAGTCCCGCAAGAGATATC 2759
Qy 2697 CGCGGCTGACAGAGTCAAGAGATGACTGAGCGGCTGCGGCACTCGCTGAGAGCGCA 2756
Db 2760 TCTG---GAGCAGAGGTCAAGAAATATCTCGGGGTGTCCGAGTCTCTGAGAGGGA 2816
Qy 2757 GGGCGGCGCAGCGGCAACAAATGATGACGCTGTGAGCAGCGGCAACAGCAGCGCTC 2816
Db 2817 TGGCCAGACCGGGCCAAACCAATGACATGACGTGTGCTGTACCGGAACACAGCGCTG 2876
Qy 2817 CATGTGATCTTCAAGATGCAAGTGAAGATCAATGTTGTGGAACCTTCTGAGAAAT 2876
Db 2877 CAGGGAGACCAAGTGAAGCTGTGCAAGAACAGCTTTCAAGATATCTCTAAGAAAGTT 2936
Qy 2877 CAAAAACAGAACGGGTGCAAGAGTGTGGGTGTGTTCAAAACTTCTGCTGTCTT 2936
Db 2937 CAAAAACAGTCAATGGCTGAGAGAGCTTGGGTGTGTTTACCAACTTCTGTTTCTT 2996

Qy 2937 CTACAAATCACACAGCAATCATCCCTTGCAGGCTGCTGTGCTACCTCGT 2996
Db 2997 CTACAAACTCATCAGATGACTACCTGCGAGGCTCCCGCTGTGGGTACAGGCT 3056
Qy 2997 CACCATCCCTCTGAGTCCGAAATCATCCAGAAAGACTGTTTCAAGCTGCACTTCA 3056
Db 3057 GAGCATCCCGAGGGAGGCGGATGGCATACAAAGACTATGTTTCAAGCTCAGTTCA 3116
Qy 3057 GTCCCACTTACTTACTTCAAGGGCGGAAGGAGTACAGCTTCAAGATGATGAAT 3116
Db 3117 ATCCCACTTACTTCTTCCGGCTGAGAGCAAGTACATTTGAAAGTGAATGAGT 3176
Qy 3117 GATCCGAGTGCACACAGCTCTGCTGCGGAGCCC 3151
Db 3177 GATCCAGGGGCGCAGACTGACCGGAGAGGCCC 3211

RESULT 10
ACLS5653
ID ACLS5653 standard; cDNA; 637 BP.
XX
AC ACLS5653;
XX
XX
DT 24-MAR-2005 (first entry)
XX
DE Human colon cancer differentially expressed polynucleotide, SEQ ID:1788.
XX
KW Differential expression; diagnosis; therapy; drug screening; cancer;
KW neoplasm; colon tumor; breast tumor; pancreas tumor; cytostatic; vaccine;
KW ss.
XX
OS Homo sapiens.
XX
PN MO200500087-A2.
PD
PD 06-JAN-2005.
PF 13-MAY-2004; 2004MO-US015421.
PR 03-JUN-2003; 2003US-0475872P.
XX
XX
PA (CHIR) CHIRON CORP.
XX
DR Randazzo F, Moler E, Escobedo J, Garcia PD;
XX
XX WPI; 2005-075421/08.
XX
XX
PT New isolated polynucleotides, which are differentially expressed in colon
PT cancer cell, useful for treating cancer, e.g. colon cancer, breast
PT cancer, or pancreatic cancer.
XX
XX
PS Claim 1; SEQ ID NO 1788; 97bp; English.

The invention relates to 9672 polynucleotides (ACLS5866-ACLS6337) which are differentially expressed in colon cancer cells. The invention also relates to vectors and host cells comprising a differentially expressed polynucleotide of the invention; a method for detecting a cancerous cell by detection of a gene product of the polynucleotides; a method for inhibiting a cancerous phenotype of a cell by inhibiting a gene product of the polynucleotides; a method of treating an individual with cancer by administration of a modulator of a gene product of the polynucleotides; and an isolated antibody that specifically binds to a polypeptide encoded by one of the 9672 polynucleotides. The polynucleotides, polypeptides, antibodies, and methods are useful for the detection of cancerous cells; for the diagnosis, prognosis and management of cancer; for the identification of agents that modulate the phenotype of cancerous cells; for the identification of therapeutic targets for cancer chemotherapy; and for the treatment of cancer, especially colon cancer and metastasized colon cancer, but also breast or pancreatic cancer. The polynucleotides are also useful as a source of probes or primers for use in diagnostic methods. The differentially expressed polynucleotides or their encoded proteins can additionally be used as vaccines to modulate primary immune

CC responses for the prevention or treatment of cancer. The present sequence
CC represents a specifically claimed polynucleotide which is differentially
CC expressed in colon cancer. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC http://wipo.int/pub/published_pcr_sequences

80 Sequence 637 BP; 144 A; 192 C; 153 G; 148 T; 0 U; 0 Other;

Query Match	18.0%	Score 620.6;	DB 14;	Length 637;
Best Local Similarity	99.4%	Pred. No. 9.1e-158;		
Matches 623;	Conservative 0;	Mismatches 4;	Indels 0;	Gaps 0;

QY	2779	ATGCTGCACGCTGCTCTGGCACCAGCAACAGCGCTCTCATGCTGGAATTACAGATCCCA	2838
Db	11	ATGCTGCACGCTGCTCTGGCAGCGGCAACACAGCGCTCTCATGCTGGAATTACAGATCCCA	70
QY	2839	GTGGAGAAATCATGTTGTCTGGAAACCTGCTGAGAAATTCAAAACAGCAAGGCTGGCAG	2898
Db	71	GTGGAGAAATCATGTTGTCTGGAAACCTGCTGAGAAATTCAAAACAGCAAGGCTGGCAG	130
QY	2899	AAGCTGAGGGGCGTGTTCACAAACTTCGCGCTGTTCTTCAATAACACAGAGCAAT	2958
Db	131	AAGCTGAGGGGCGTGTTCACAAACTTCGCGCTGTTCTTCAATAACACAGAGCAAT	190
QY	2959	CATCCCTGTGCGACGCTGCTCTGTCTCGCTACTCGCTACACATCCCTCGAGCTCGAG	3018
Db	191	CATCCCTGTGCGACGCTGCTCTGTCTCGCTACTCGCTACACATCCCTCGAGCTCGAG	250
QY	3019	AACATCCAGAAAGACTACGTGTTTCAAGCTGCACTTCAAGTCCACGCTTACTTACG	3078
Db	251	AACATCCAGAAAGACTACGTGTTTCAAGCTGCACTTCAAGTCCACGCTTACTTACG	310
QY	3079	GCGGAAAGCGAGTACAGTTGGAAGGGAATGGAAGATCCGAGGACCAAGCTCT	3138
Db	311	GCGGAAAGCGAGTACAGTTGGAAGGGAATGGAAGATCCGAGGACCAAGCTCT	370
QY	3139	GCCTCGGACCCCACTGTTGAGCCACAAAGACTCTTGTGTATTGATGCGCGGACACA	3198
Db	371	GCCTCGGACCCCACTGTTGAGCCACAAAGACTCTTGTGTATTGATGCGCGGACACA	430
QY	3199	CTGCTTTCCGCACTGCTGCTTCTCTGGAAGACTTCTTCTTCTGTATTATGAAGC	3258
Db	431	CTGCTTTCCGCACTGCTGCTTCTCTGGAAGACTTCTTCTTCTGTATTATGAAGC	490
QY	3259	CTGCTAAATTAACACCTGCTGAAATTCAAAAATGAGGCTTCCAGAGCTCTCTGTC	3318
Db	491	CTGCTAAATTAACACCTGCTGAAATTCAAAAATGAGGCTTCCAGAGCTCTCTGTC	550
QY	3319	TCCACAGCCGCGTTTTTAAACCCGACCTCTCAGCGTTTGATGAAACAGCGCTCCCACT	3378
Db	551	TCCACAGCCGCGTTTTTAAACCCGACCTCTCAGCGTTTGATGAAATGAGCAGCGCTCCCACT	610
QY	3379	CCAAGTCTTGCACTCCGCTGGGGGGGCT	3405
Db	611	CCAAGTCTTGCACTCCGCTGGGGGGGCT	637

RESULT 11
ACD92278/c
ID ACD92278 standard; cDNA; 582 BP.

AC ACD92278;

DT 23-SEP-2003 (first entry)

DE Human colon cancer cell expressed cDNA #690.

KM Open reading frame detection; genome sequencing; colon cancer;
KM breast cancer; population genome analysis; genetic shift; cancer;
KM antibiotic resistance; antibiotic non-tolerance; congenital disease;
KM agriculture; food crop genome; resistance gene; retrovirus;
KM influenza virus; eukaryotic pathogen detection; trypanosome; Plasmodium

KW gene; SS

OS Homo sapiens.

PN US2002155438-A1

PD 24-OCT-2002

PF 27-SEP-1999; 99US-00406117.

PR 20-NOV-1998; 98US-00196716.

PA (SIMP/) SIMPSON A J G.

PA (BREN/) BRENTANI R R.

PI Simpson AJG, Neto ED, Brentani RR;

DR WPI; 2003-182626/18.

PT Determining open res

PT Determining open reading frames of genome of an organism e.g. a human
PT suffering from cancer involves use of single oligonucleotide primer at
PT low stringency for preparing single-stranded cDNA from mRNA of
PT individual.

PS Example 9; Page 105; 959pp; English

The invention describes a method for determining open reading frames in the genome of organism, comprising contacting mRNA from cell of organism with a single oligonucleotide primer (I) at low stringency, preparing single-stranded cDNA by reverse transcribing mRNA with (I), amplifying cDNA, sequencing the product, and repeating the contacting, preparing and amplifying steps with different primers and sequencing resulting nucleic acids. The method is useful for: determining that a known nucleotide sequence of a genome of an organism corresponds to a nucleotide sequence of an open reading frame; for preparing a contig, nucleic acid molecule from a genome of an organism; and for sequencing all or part of a genome of an organism. mRNA is obtained from mammalian or human cell which is associated with a pathological condition e.g. a colon cancer or breast cancer cell. The method is useful for analyses of populations of subjects and can be used to carry out genetic analyses of large or small populations. Further, it can be used to study living systems to determine if, e.g. there have been genetic shifts which render an individual or population more or less likely to be afflicted with diseases such as cancer, to determine antibiotic resistance or non-tolerance, and so forth. The method can also be used in the study of congenital diseases, and the risk of affliction to a fetus, as well as the study of whether the conditions are likely to be passed to offspring through ova or sperm. The analyses for pathological conditions can be carried out in all animals, plants, birds, fish, etc. Using this method, in the area of agriculture, for example the genomes of food crops can be studied to determine if resistance genes are present, defects in plant genomes can also be studied in this way. Similarly, the method permits determination of the pathogens which integrate into the genome, such as retroviruses and other integrating viruses such as influenza virus, have undergone shifts or mutations, which may require different approaches to therapy. This method is also applied to eukaryotic pathogens, such as trypanosomes, different types of Plasmodium, etc. The method essentially eliminates sequencing of non-coding portions. This sequence represents a polynucleotide isolated from human colon cancer cell cDNA library

Sequence 582 BP; 109 A; 147 C; 185 G; 140 T; 0 U; 1 Other;

Query Match	16.1%	Score	554.2	DB	10	length	582
Best Local Similarity	99.3%	Pred. No.	1e-139				
Matches	556	Conservative	0	Mismatches	4	Indels	0
						Gaps	0

QY 2631 CACGAGCCCCCGCCCTGACGATTCCTGGCCAGAGCCCCCTGACAAACAAGTCCCTGATGA 2630

Db 571 CAGCAGCCCCCGCCCTGAGTTCCTGGCCAGAGCCCCCTGACAAACAAGTCCCTGATGA 512

QY 2691 AGCCACCGCGGCTGACCTAGAGATCAGAGATGACTGAGGCGCTTCGGGACATGCGCTGGA 2750

Db	511	AGCCACCGCGGCTGAACAGAGTACAGAGNATGACTTGAGGCCCTTCGGCACAATGCGCTGGA	452
Qy	2751	GCGCACGAGCGCCCGGACCGCGGCAACACAACTGGTGACCGTGTGCTGGCACCGCAACACCG	2810
Db	451	GCGCACGAGCGCCCGGACCGCGGCAACACAACTGGTGACCGTGTGCTGGCACCGCAACACCG	392
Qy	2811	CGTCTCCATGGATGGACTTCAGCATTCGCGAGTGGAGAAATAGTTGTGCTGAAACCTGCTGAG	2870
Db	391	CGTCTCCATGGATGGACTTCAGCATTCGCGAGTGGAGAAATAGTTGTGCTGAAACCTGCTGAG	332
Qy	2871	GAATATCAAAAACAGCAACGGGTGGCAGAAAGCTGTGGTGTGTTACAAACCTTCTGCTT	2930
Db	331	GAATATCAAAAACAGCAACGGGTGGCAGAAAGCTGTGGTGTGTTACAAACCTTCTGCTT	272
Qy	2931	GTTCTTTCACAAATACACACAGGACAAATCATCCCTTCGACAGCCGTGCTGCTGCGCTA	2990
Db	271	GTTCTTTCACAAATACACACAGGACAAATCATCCCTTCGACAGCCGTGCTGCTGCGCTA	212
Qy	2991	CTCGCTCACATCCCTCTGTGATCCGAGAACATCCAGAAAGCTACGTTCAAGCTGCA	3050
Db	211	CTCGCTCACATCCCTCTGTGATCCGAGAACATCCAGAAAGCTACGTTCAAGCTGCA	152
Qy	3051	CTTCAAGTCCCAAGTCTACTTCACTTCAGGGCGGAAAGCAAGTACAGCTTCGAAAGTGGAT	3110
Db	151	CTTCAAGTCCCAAGTCTACTTCACTTCAGGGCGGAAAGCAAGTACAGCTTCGAAAGTGGAT	92
Qy	3111	GGAAGTAGTCCGCAAGTGCACCAAGCTCTGCTCGGACCCCAAGTGTGAGCCACAAAGA	3170
Db	91	GGAAGTAGTCCGCAAGTGCACCAAGCTCTGCTCGGACCCCAAGTGTGAGCCACAAAGA	32
Qy	3171	GTTCTTGTGTATTTGATGGC	3190
Db	31	GTTCTTGTGTATTTGAAAGAC	12

RESULT 12
ACD94832/c
ID ACD94832 standard; cDNA; 582 BP.
XX AC
XX ACD94832;
XX
DT 23-SEP-2003 (first entry)
XX
DE Human colon cancer cell expressed cDNA #3244.
XX
KW Open reading frame detection; genome sequencing; colon cancer;
KW breast cancer; population genome analysis; genetic shift; cancer;
KW antibiotic resistance; antibiotic non-tolerance; congenital disease;
KW agriculture; food crop genome; resistance gene; retrovirus;
KW influenza virus; eukaryotic pathogen detection; trypanosome; Plasmodium;
gene; ss.
XX
XX Homo sapiens.
OS
XX US2002155438-A1.
PN
XX 24-OCT-2002.
PD
XX
PF 27-SEP-1999; 99US-00406117.
XX
PR 20-NOV-1998; 98US-00196716.
XX
XX (SIMP/) SIMPSON A J G.
PA (NETO/) NETO E D.
PA (BREN/) BRENTANI R R.
XX
PI Simpson AJG, Neto ED, Brentani RR;
XX
DR WPI; 2003-182626/18.
XX
PT Determining open reading frames of genome of an organism e.g. a human
PT suffering from cancer involves use of single oligonucleotide primer at
PT low stringency for preparing single-stranded cDNA from mRNA of

PT individual.

Example 9; Page 477; 959pp; English

The invention describes a method of determining open reading frames in the genome of organism, comprising contacting mRNA from cell of organism with a single oligonucleotide primer (I) at low stringency, preparing single-stranded cDNA by reverse transcribing mRNA with (I), amplifying cDNA , sequencing the product, and repeating the contacting, preparing and amplifying steps with different primers and sequencing resulting nucleic acids. The method is useful for: determining that a known nucleotide sequence from a genome of an organism corresponds to a nucleotide sequence of an open reading frame; for preparing a contig, nucleic acid molecule from a genome of an organism; and for sequencing all or part of a genome of an organism. mRNA is obtained from mammalian or human cell which is associated with a pathological condition e.g. a colon cancer or breast cancer cell. The method is useful for analyses of populations of subjects and can be used to carry out genetic analyses of large or small populations. further, it can be used to study living systems to determine if, e.g. there have been genetic shifts which render an individual or population more or less likely to be afflicted with diseases such as cancer, to determine antibiotic resistance or non-tolerance, and so forth. The method can also be used in the study of congenital diseases, and the risk of affliction to a foetus, as well as the study of whether the conditions are likely to be passed to offspring through ova or sperm. The analyses for pathological conditions can be carried out in all animals, plants, birds, fish, etc. Using this method, in the area of agriculture, for example the genomes of food crops can be studied to determine if resistance genes are present, defects in plant genomes can also be studied in this way. Similarly, the method permits determination of the pathogens which integrate into the genome, such as retroviruses and other integrating viruses such as influenza virus, have undergone shifts or mutations, which may require different approaches to therapy. This method is also applied to eukaryotic pathogens, such as trypanosomes, different types of Plasmodium, etc. The method essentially eliminates sequencing of non-coding portions. This sequence represents a polynucleotide isolated from human colon cancer cell cDNA library

Sequence 582 BP; 109 A; 147 C; 185 G; 140 T; 0 U; 1 Other;

Query Match 16.1%; Score 554.2; DB 10; Length 582;

```

best local similarity  55.58; 1200; 1200;
Matches 556; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

QY	2631	CAGCAGCCCGCCCTGAGTTCCTGGCCAGCAGCCCCCTGACAAACAAGTCCCTGATGA	2650
Db	571	CAGCAGCCCCCCTTGATGTTCTGGCCAGCAGCCCCCTGACAAACAAGTCCCTGATGA	512
QY	2631	AGCCACCGGGCTGACCCAGAGTCAAGAGTGAACCTTAGAGCCCTCGGCACATCGCTGGA	2750
Db	511	AGCCACCGCGCTGACCGAGGTAGAGNATGACTAGAGCCCTCGGCACATCGCTGGA	452
QY	2751	GCGCCAGGCCCCCGCAGCCGCGCAACACATGTGTCACGTGTGCTGGCACCGCAACACCG	2810
Db	451	GCGCCAGGCCCCCGCAGCCGCGCAACACATGTGTCACGTGTGCTGGCACCGCAACACCG	392
QY	2811	CGTCCCATGGTGGACTTCAGCATCCGACGTGGAGAAATCAGTTGTCTGGAAACTGCTAG	2870
Db	391	CGTCCCATGGTGGACTTCAGCATCCGACGTGGAGAAATCAGTTGTCTGGAAACTGCTAG	332
QY	2871	GAATATCAAAAACAGCAACGGGTGGCAGAACTGTGGTGTGTTCACAAACTTCTGCT	2930
Db	331	GAATATCAAAAACAGCAACGGGTGGCAGAACTGTGGTGTGTTCACAAACTTCTGCT	272
QY	2931	GTTCTTCAACAATCACACAGACAATCATCCCTTTCAGCCTGCTCTGCTCGGCTA	2990
Db	271	GTTCTTCAACAATCACACAGACAATCATCCCTTTCAGCCTGCTCTGCTCGGCTA	212
QY	2991	CTCGCTCAACATCCCTCTGAGTCCGAGAACATCCAGAAAGCTACGTTCAAGCTGCA	3050
Db	211	CTCGCTCAACATCCCTCTGAGTCCGAGAACATCCAGAAAGCTACGTTCAAGCTGCA	152
QY	3051	CTTCAGTCCACGCTCTACTACTTCAGGCGGAAAACGAGTACACGTTTCGAAAGGTGAT	3110

Db	151	CTTCAAGTCCACGCTCTACTTCTCAGGCGGAAAGCAGTACAGCTTCGAAAGTGGAT	92
Qy	3111	GGAAAGTATCCGCAAGTGGCCACCAAGCTGTGCTCGGACCCCAAGTGTGAGCCACAAGA	3170
Db	91	GGAAAGTATCCGCAAGTGGCCACCAAGCTGTGCTCGGACCCCAAGTGTGAGCCACAAGA	32
Qy	3171	GTCCTCTGTGATTTGATGGC	3190
Db	31	GTCCTCTGTGATTTGATGAAC	12
RESULT 13			
ACN41157/C	ID	ACN41157 standard; cDNA; 6503 BP.	
AC	ACN41157;		
XX	18-NOV-2004	(first entry)	
DT			
XX			
DS		Human diagnostic and therapeutic polynucleotide SEQ ID NO:32.	
XX			
KW	ss; gene; gene therapy; human diagnostic and therapeutic polynucleotide;		
XX	dthp.		
OS	Homo sapiens.		
XX	MO2004023973-A2.		
PN			
XX	25-MAR-2004.		
PD			
PF	12-SEP-2003; 2003WO-US028227.		
XX			
XX	12-SEP-2002; 2002US-0410259P.		
PR	12-SEP-2002; 2002US-0410260P.		
XX			
PA	(INCY-) INCYTE CORP.		
PI	Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F;		
PI	Hatchborne TA, Suchorolski MT, Altus CM, Pitesa SJ, Elder LV;		
PI	Mooney EM, Delegeane AM, Panesar IS, Banville SC, Reddy TP;		
PI	Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstin EH;		
PI	Peralta CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LJ;		
PI	Legace RE, Spito PA, Stewart EA, Wengrove J, Vite UA, Kitson ES;		
PI	Xu Y, Khong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D;		
PI	Patury S, Shi X, Suarez CJ;		
XX			
DR	WPI: 2004-329368/30.		
XX	P-PsDB; ABM62505.		
XX			
PT	New diagnostic and therapeutic polynucleotides and polypeptides, useful		
PT	in diagnosing a condition, disease or disorder associated with human		
PT	molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or		
PT	in gene mapping.		
XX			
PS	Claim 1, Page; 190pp; English.		
CC			
CC	The invention relates to novel diagnostic and therapeutic polynucleotides		
CC	selected from one of the 2722 sequences defined in the specification. A		
CC	polynucleotide of the invention may have a use in gene therapy. The human		
CC	diagnostic and therapeutic polynucleotides (dthp) or polypeptides may be		
CC	used to diagnose a particular condition, disease or disorder associated		
CC	with human molecules, e.g. cell proliferative disorders,		
CC	autoimmune/inflammatory disorder, developmental disorder, endocrine		
CC	disorder, neurological disorder, gastrointestinal disorders, or		
CC	infections caused by virus, bacteria, fungi or parasite. The dthp		
CC	molecules may also be used in genetic mapping, in identifying individuals		
CC	from minute biological samples, in detecting single nucleotide		
CC	polymorphisms, as molecular weight markers, and for somatic or germline		
CC	gene therapy. The present sequence represents a dthp polynucleotide of		
CC	the invention. Note: The sequence data for this patent is not represented		
CC	in the printed specification, but was obtained in electronic format		
CC	directly from WFO at www.wipo.int/pct/en/sequences/lilectn.htm		

Seq	Sequence	6503 BP, 1670 A, 1409 C, 1682 G, 1742 T, 0 U, 0 Other;
XX		
XX	Query Match	14.3%; Score 490.8; DB 13; Length 6503;
XX	Best Local Similarity	99.6%; Pred. No. 6.6e-122;
XX	Matches 492; Conservative	0; Mismatches 2; Indels 0; Gaps 0;
QY	2949	CCAGGACATATATCCCTTGGCAGCTGCTCTGCTGGCTACTGCTCACCATCCCTC 3008
DB	6435	CCAGGACATATATCCCTTGGCAGCTGCTCTGCTGGCTACTGCTCACCATCCCTC 6376
QY	3009	TGAGTCGAGAACTCCGAAAGACTACAGTTCAGCTGACCTTCAAGTCCACGCTCA 3068
DB	6375	TGAGTCGAGAACTCCGAAAGACTACAGTTCAGCTGACCTTCAAGTCCACGCTCA 6316
QY	3069	CTACTTCGAGGCGGAAAGCGAGTACACGTTTGAAAGGTGATGGAAGTATCCGACATGC 3128
DB	6315	CTACTTCGAGGCGGAAAGCGAGTACACGTTTGAAAGGTGATGGAAGTATCCGACATGC 6256
QY	3129	CACCAAGCTCTGCTTCGCGGACCCCAAGTGTGAGCCACAAAGAGTCTTGTGTATGATG 3188
DB	6255	CACCAAGCTCTGCTTCGCGGACCCCAAGTGTGAGCCACAAAGAGTCTTGTGTATGATG 6196
QY	3189	GCCGAGACACATCCGTTTCCGAGGAGGCTGCTTCTCGAAGACGTTTCCCTTCTTGTA 3248
DB	6195	GCCGAGACACATCCGTTTCCGAGGAGGCTGCTTCTCGAAGACGTTTCCCTTCTTGTA 6136
QY	3249	TTAATGAAGCTGTGTAATAATTAACACCTGTCTGAAATCAAAAACATGAGCTCCACAG 3308
DB	6135	TTAATGAAGCTGTGTAATAATTAACACCTGTCTGAAATCAAAAACATGAGCTCCACAG 6076
QY	3309	CTCTCTGTCTCCACAGCGCGGTTTTTAAACCCGACCTCTCAGCGTTTGATGAACAGC 3368
DB	6075	CTCTCTGTCTCCACAGCGCGGTTTTTAAACCCGACCTCTCAGCGTTTGATGAACAGC 6016
QY	3369	GCTCCCACTTCGATCCTTGAGATCCGCTGGGGGGGCTGTTCTTAACTAGTGCAGATAT 3428
DB	6015	GCTCCCACTTCGATCCTTGAGATCCGCTGGGGGGGCTGTTCTTAACTAGTGCAGATAT 5956
QY	3429	AAACATTGTGATT 3442
DB	5955	AAACATTGTGATT 5942
XX		
XX	RESULT 14	
XX	ABL67912/c	
XX	ID ABL67912	standard; DNA; 506 BP.
XX	AC ABL67912;	
XX	DT	
XX	15-MAY-2002	(first entry)
XX	DE	
XX	Ovary cancer related gene sequence SEQ ID NO:6249.	
XX	Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;	
XX	stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;	
XX	cytostatic; gene therapy; antineoplastic; Wilms' tumour; adenocarcinoma;	
XX	gene; ds.	
XX	OS	
XX	Homo sapiens.	
XX	XX	
XX	WO200194629-A2.	
XX	PD	
XX	13-DEC-2001.	
XX	PF	
XX	30-MAY-2001; 2001WO-US010838.	
XX	05-JUN-2000; 2000US-0209473P.	
XX	PR 05-JUN-2000; 2000US-0209531P.	
XX	PR 18-SEP-2000; 2000US-0233133P.	
XX	PR 18-SEP-2000; 2000US-0233617P.	
XX	PR 20-SEP-2000; 2000US-0234009P.	
XX	PR 20-SEP-2000; 2000US-0234034P.	

20-SEP-2000; 2000US-02345052P.
PR 22-SEP-2000; 2000US-0234509P.
PR 22-SEP-2000; 2000US-0234567P.
PR 25-SEP-2000; 2000US-0234923P.
PR 25-SEP-2000; 2000US-0234924P.
PR 25-SEP-2000; 2000US-0235077P.
PR 25-SEP-2000; 2000US-0235082P.
PR 25-SEP-2000; 2000US-0235134P.
PR 25-SEP-2000; 2000US-0235280P.
PR 26-SEP-2000; 2000US-0235637P.
PR 26-SEP-2000; 2000US-0235638P.
PR 27-SEP-2000; 2000US-0235711P.
PR 27-SEP-2000; 2000US-0235720P.
PR 27-SEP-2000; 2000US-0235840P.
PR 27-SEP-2000; 2000US-0235863P.
PR 28-SEP-2000; 2000US-0236028P.
PR 28-SEP-2000; 2000US-0236032P.
PR 28-SEP-2000; 2000US-0236033P.
PR 28-SEP-2000; 2000US-0236034P.
PR 28-SEP-2000; 2000US-0236109P.
PR 28-SEP-2000; 2000US-0236111P.
PR 29-SEP-2000; 2000US-0236842P.
PR 29-SEP-2000; 2000US-0236891P.
PR 02-OCT-2000; 2000US-0237172P.
PR 02-OCT-2000; 2000US-0237173P.
PR 02-OCT-2000; 2000US-0237278P.
PR 02-OCT-2000; 2000US-0237294P.
PR 02-OCT-2000; 2000US-0237295P.
PR 02-OCT-2000; 2000US-0237316P.
PR 03-OCT-2000; 2000US-0237425P.
PR 03-OCT-2000; 2000US-0237598P.
PR 03-OCT-2000; 2000US-0237604P.
PR 03-OCT-2000; 2000US-0237606P.
PR 03-OCT-2000; 2000US-0237608P.
PR 03-OCT-2000; 2000US-0237608P.
PR 01-NOV-2000; 2000US-0248467P.
PR 01-NOV-2000; 2000US-0245084P.

(AVAL-) AVALON PHARM.
XX
PA Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI Soppe DR, Weaver Z;
FI
DR WPI; 2002-188264/24.
XX
XX
PT Screening for anti-neoplastic agent involves exposing cells to a chemical
PT agent to be tested for anti-neoplastic activity, and determining a change
PT in expression of a gene of a signature gene set.
XX
XX
PS Claim 1; SEQ ID NO 6249; 44pp; English.
XX
XX The present invention describes a method (M1) for screening for an anti-
CC neoplastic agent. The method involves exposing cells to a chemical agent
CC to be tested for anti-neoplastic activity, determining a change in (I)
CC expression of at least one gene (I) of a signature gene set, where (I)
CC comprises a sequence (S) selected from 8447 sequences (given in ABL51664
CC to ABL70110), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (I) has cytosolic
CC activity and can be used in gene therapy. M1 can be used for screening an
CC anti-neoplastic agent, and can be used for producing a product which is
CC the data collected with respect to the anti-neoplastic agent as a result
CC of M1, and the data is sufficient to convey the chemical structure and/or
CC properties of the agent. M1 can be used in the treatment of cancer such
CC as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,
CC prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell
CC cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous
CC cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilms'
CC tumour.

Sequence 506 BP; 118 A; 112 C; 135 G; 141 T; 0 U; 0 Other;

Query Match	13.8%	Score 474.4	DB 6	Length 506
Best Local Similarity	99.6%	Pred. No. 4.7e-118		
Matches 486; Conservative	0	Mismatches	1	Gaps 1

QY	2824	GACCTTCAGCATCCGACGTGGAGAAATCACTGTCTCTGAAAACCTGCTGAGGAAATTCAAAAC	2883
Db	506	GACTTCAGCATCCGACGTGGAGAAATCACTGTCTCTGAAAACCTGCTGAGGAAATTCAAAAC	447
QY	2884	AGCAACGGGTGGCAGAAAGCTGTGGGTGTGTTCACAAACCTCTGCTCTTCTACAAA	2943
Db	446	AGCAACGGGTGGCAGAAAGCTGTGGGTGTGTTCACAAACCTCTGCTCTTCTACAAA	387
QY	2944	TCACACCAAGACATTCATCCCTCTGCGACGCTGCTCTGCTCGGCTACTGCTCAACATC	3003
Db	386	TCACACCAAGACATTCATCCCTCTGCGACGCTGCTCTGCTCGGCTACTGCTCAACATC	327
QY	3004	CCCTCTGAGTCCGAGAACATCCAGAAAGACTACGTGTTCAAGCTGCACTTCAATCCAC	3063
Db	326	CCCTCTGAGTCCGAGAACATCCAGAAAGACTACGTGTTCAAGCTGCACTTCAATCCAC	267
QY	3064	GTCATCTACTTCAGGGCGGAAAGCGAGTACAGTTTCGAAAGGTGATGGAAGTATCCG	3123
Db	266	GTCATCTACTTCAGGGCGGAAAGCGAGTACAGTTTCGAAAGGTGATGGAAGTATCCG	207
QY	3124	AGTGCACACAGCTCTGCTCGCGAACCCACGCTGTGAGCCACAAAGAGCTCTTGTGTAT	3183
Db	206	AGTGCACACAGCTCTGCTCGCGAACCCACGCTGTGAGCTACAAAGAGCTCTTGTGTAT	147
QY	3184	TGATGGCCGAGACACTCTGTTCCGACAGTGCTGTTCTGGAAGACGTTTCTTTCTT	3243
Db	146	TGATGGCCGAGACACTCTGTTCCGACAGTGCTGTTCTGGAAGACGTTTCTTTCTT	87
QY	3244	CTGATTTAATAAACCTGTGTAATTTAAACCTGTCTGAAAAATCAAAAACATGGCTTCCC	3303
Db	86	CTGATTTAATAAACCTGTGTAATTTAAACCTGTCTGAAAAATCAAAAACATGG-TTCCC	28
QY	3304	AGCAGCTC 3311	
Db	27	AGCAGCTC 20	

```

RESULT 15
ADQ63863
ID ADQ63863 standard; cDNA, 3200 BP.
XX
XX ADQ63863;
AC
XX
XX 07-OCT-2004 (first entry)
DT
XX
XX Novel human cDNA sequence #1024.
DE
XX
XX ss; gene; osteopathic; neuroprotective; nootropic; antiparkinsonian;
KW cyostatic; gene therapy; diagnostic marker; morbid state; osteoporosis;
KW neurological disease; Alzheimer's disease; Parkinson's disease; dementia
KW cancer.
XX
XX Homo sapiens.
OS
XX
XX EPI440981-A2.
EN
XX
XX 28-JUL-2004.
PD
XX
XX 21-JAN-2004; 2004EP-00001196.
PF
XX
XX 21-JAN-2003; 2003JP-00102206.
PR
XX
XX 09-MAY-2003; 2003JP-00131392.
PR
XX
XX (REAS-) RES ASSOC BIOTECHNOLOGY.
PA
XX
XX Isegaki T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Iehi S;
PI Yamamoto J, Isono Y, Nagai K, Irie R;
PI WPI; 2004-535376/52.
DR
XX
XX P-PSDB; ADQ66051.
XX
XX Novel 2495 cDNA, useful for treating osteoporosis, neurological diseases,

```

PT Alzheimer's diseases, Parkinson's diseases, dementia and various cancers.
 XX
 PS Claim 1; SEQ ID NO 1024; 2449bp; English.

CC The invention relates to 2495 novel polynucleotides (I) and their encoded
 CC polypeptides, sequences hybridizing to these nucleotides, sequences
 CC encoding partial polypeptides and sequences having 70% or 90% identity to
 CC the nucleotide and protein sequences. The nucleotides and polypeptides
 CC are useful as diagnostic markers or therapeutic target for the diseases
 CC or morbid states. They are also useful for treating osteoporosis,
 CC neurological diseases, Alzheimer's diseases, Parkinson's diseases,
 CC dementia and various cancers. This sequence corresponds to a nucleotide
 CC sequence of the invention.

XX Sequence 3200 BP; 951 A; 659 C; 718 G; 872 T; 0 U; 0 Other;

Query Match 13.0%; Score 447; DB 12; Length 3200;

Best Local Similarity 65.5%; Pred. No. 3.7e-110;

Matches 654; Conservative 0; Mismatches 345; Indels 0; Gaps 0;

QY 144 CCCAACCCTTCAGAAAACCTCGTCATCAAAATCCAGATGTGATACACCCAGCA 203
 DB 160 CCCAAGGCGCTCGCTGAGATGTACATTTAAAGTCAGATTTTGGATGATCCAGAA 219
 QY 204 GGCATTGAAAGTCCAAAGAGCTCTGGGAAAGTGCTGTGATGCAAGTTGCAACCA 263
 DB 220 GATTTTGTGTTGATCAAAAGTCATCCGGAAAGCAATTTTAACTGAGTTGCAACCA 279
 QY 264 CCTCAACCTCGTGAAGTGATATTTTGGCTCGAGTTTCTGATCAGAAAAGATCAC 323
 DB 280 TCTAAATCTTGCGAAGAAAGATATTTTGGATTGAAATTCGACGATTCGGAATATA 339
 QY 324 GGTGTGCTGATCTCTTAAACCATTTGAAACAGATTGAAGGCCAAAGCAGTTGT 383
 DB 340 TGTTTGGCTGAGCTTTTGAAAGCCATTAACAAAGCAGTAAATAATCTTAAGAGATGT 399
 QY 384 TGTTAAGTTTGTGTAATTTCTTCCGCTGACCAACACCACTCAAGAAAGACTCAC 443
 DB 400 TTTCAATTTATGTGTAATTTTCCAGTGAACCTGAGCATCTGGGAAAGAACTTAC 459
 QY 444 AAGTACCTGTTCGCTGAGGTGAAGAGCACTTGGCTCAAGGCAAGTTGACGTGTAA 503
 DB 460 AAGGTATCTTTTACTTCAATAAAGAGATTGGCTTAGAAGGCTTCATGACAG 519
 QY 504 TGACACGAGCGAGCTCTCTTGAATTCACATTTGCAATCTGAGATTGGGATTTTGA 563
 DB 520 TGACACCTGTACGCGTTGATGTATCTCAATCTTACATCAGAACTTGAGACTTTCA 579
 QY 564 TGAAGCTTGACAGAGAGCACTTAGCAAAATAATATACACTCAGCAGAGCGACT 623
 DB 580 TGAAGAAACAGATAGAGAGCATCTGGCAAACTCGTACTTACAAACCAAGACTGTTT 639
 QY 624 AAGAGCAAAATGTGGAATTTTCACATTAACAATTGGAACAAACACAGCAGATCAGA 683
 DB 640 AAGAGGCAAGATCATGCACTTTCATCAGAAAGCAATTGGCAGAGCCAGCTGAATCTGA 699
 QY 684 TTTCCAGCTCTTGAAGATTGCCGCTGAGATGTATGAAATCCGTTGCAACCCGCG 743
 DB 700 CATTCGCTACTGAGCATAGCAAGAGCTGATATGTATGTCATCAGGCTTCAACCCCG 759
 QY 744 CAAGGACAGGAAAGGACAGAGATCATCTGCGCTTGCCACACGAGGAATTTAGTGT 803
 DB 760 CAGTATGTGTAAGAGATGCAAGTTTCACTGCGCTGTTGCTCACAATGGAGTACTGTT 819
 QY 804 TCAGGCTTCACTAAGATCAATGCTTCAACTGGGCCAAGGTGCGGAAGCTGAGCTTCAA 863
 DB 820 ACGGGAAATACAAAGATCAATCTTTAACTGGGCTAAAAATCCGCAAGTTGAGTTTAA 879
 QY 864 GAGGAAGCGTTTCTCATCAAGCTCCGGCCAGATGCAATAGTCGTAACAGATACCTT 923
 DB 880 GAGAAAGCATTTTCTCATCAAACTTCAATGCAATATCTTGTGTGCAAGGATACCTT 939
 QY 924 GGAATTCGTATGCGCAGTTCGGAATTTCTGCAAGTCTTCTGGAATAATCTGTGTGAACA 983

DB 940 GAGATTCACATGCGCAGCGAGATGCTCTGCAAGCTTTCTGGAAGACTTGTGTGAATA 999
 QY 984 TCATGCTCTTCTTGAAGCTTTTGAAGAGCCCAAAAGCCCAAGCCCGCTCTTTAG 1043
 DB 1000 CCATGCTTTCTTCAAGCTTTTGGAAAGAGCCCAATCAAGCCCAAGCCCTACTCTGCAG 1059
 QY 1044 CCGGGGTCATATCTTGGTTGAGTGTGAGTCTGAAAGCAGGTTCTGACTATGTAA 1103
 DB 1060 CAAGGTTCAAGTTTCCGCTATGTGACGAACCAAGGCACTTTTGGATATGGAG 1119
 QY 1104 AAGAGGAGACATTAAGAGGTGCAAGTTTGAAGGAAGCA 1142
 DB 1120 AAAAGGAGGCTGAAAGAGCTTGCATTTGAAGGAACA 1158

Search completed: May 22, 2006, 01:14:34
 Job time : 1358.07 secs

This Page Blank (uspic)


```
gene
1. .475
/gene="DKFZp469c092"
100. .3237
/gene="DKFZp469c092"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAH91292.1"
/db_xref="GI:5579111"
/translation="MEIQRPTPGSRILGAPENSGISTLERGOKPPRPSGKLIVSIKI
QMLDTPQAEVFPQAPRGVLDVAVCNHLNLVBDYFGIECDHKKITTWDLRLPV
KQIRRKHVAVKVPFPPDHQQLQEBLTRYLFAIQVQDLAGQLTNDLSALLI
SHIVSEIDFPDALREHLANKNYIQDDALDEKIVEHNHNIQGTPESEPOLEI
ARLEWYGIHLPAKDRSEGTINIAVANALIVFOGTINIAFNMAKYKSLFKRF
LTKRPDANSKODTLEFLMAKSPDFCKFMKICVEHNAFPLPEERKPEKLSRSG
SEFRSGRKQOVLDYVKEGKHQUPREKSKHISIRLTAQPELSEVLEEOQOS
ASLTREGBASFEQOSCOQKPEKVSPPGSHPSVPFRSRANKQADGASALTEB
EEEVKDRTPQSKPPQPPSTSLTGSPLSELVNSQGVAPANTYLSPLSPOLKE
ASPLISPLNDQACPTDDEDEGRRRPPTDKAYFLAKESVTERTYLKDLEVTISW
QAVSEKDEMPALKSLIPNPEPLHKFTNPLKEIEORLALMEGRSNQIERYDQIG
DVALNKIOGMKHLAVHLMGSEALBENGIKSSRLEBFCDEPLOKVCYLPPLFL
LRLRHMHYKQVLERLCKHPPSHADPDCRALAGITEMVALHGMITKMEHQKL
HELKDLIGIDNLVYPRGEPFIRPLGSLSKSLSGKLOQRMFLRNDILLTSRSLTASNO
FEVHQQLPLYGMTIKESDEDEWVPHCLTLRGQROSTIVASSRSEKWEVEDIQALID
IAEKNSLAPERFLASSPPDNKSPDEATADQSEEDLDSASRTSLERQAFHRENTVHV
CHRRNTSVMDVFSVAVENQLSGNLIRKFKNSNGQKLMVVTNFCLEFFYKSHQDNHP
LASLPLGYSLTIPTESENIHKDYVKLHFKSHVYVYFRAESEYTERMEVIRSATSS
AARVIVSHKESLVY"
```

ORIGIN

```
Query Match      97.0%; Score 3339.8; DB 6; Length 4775;
Best Local Similarity 98.3%; Pred. No. 0;
Matches 3305; Conservative 0; Mismatches 57; Indels 1; Gaps 1;
```

```
OY 1 CGCCGAGCCGCGGCGCTGTGAGATATTTCTTAAGCCGCTTTCATCTGGAAGAAATA 60
DB |||
OY 52 CGCCGAGCCGCGGCGCTGTGAGATATTTCTTAAGCCATTTCACTGGAGAATAA 111
DB |||
OY 61 GAGCAGAGCCGAGCCCGCAGATCAAGACTGGGGGCCCCGGAATAATTCGGGAGTCACTACC 120
DB |||
OY 112 GAGCAGAGCCGAGCCCGCAGATCAAGACTGGGGGCCCCGGAATAATTCGGGAGTCACTACC 171
DB |||
OY 121 TTGGAACTGTGACAGAAAGCCGCCCCCAACACTTCAGAAAACTCGTGTTCATCAAAATC 180
DB |||
OY 172 TTGGAACTGTGACAGAAAGCCGCCCCCAACACTTCAGAAAACTCGTGTTCATCAAAATC 231
DB |||
OY 181 CAGATCTGTGATGACACCCAGAGAGGATTTGAAGTTCCAAAGACTCTCTGGGAAGTG 240
DB |||
OY 232 CAGATCTGTGATGACACCCAGAGAGGATTTGAAGTTCCAAAGACTCTCTGGGAAGTG 291
DB |||
OY 241 CTGCTGATGATGCAACCACTCAACCTGTGGAAGTGACTAATTTGGCTCGAG 300
DB |||
OY 292 CTGCTGATGATGCAACCACTCAACCTGTGGAAGTGACTAATTTGGCTCGAG 351
DB |||
OY 301 TTTCCTGATCAGAAAAAGATCAGGCTGTGGCTGGATCTCTTAAACCAATGTGAAACAG 360
DB |||
OY 352 TGCTCTGATCAGAAAAAGATCAGGCTGTGGCTGGATCTCTTAAACCAATGTGAAACAG 411
DB |||
OY 361 ATTTGAAGCCCAAGACCTGTGTTTAACTTTGTGTGAAAATCTTTCCGCTGCAAC 420
DB |||
OY 412 ATCAGAAAGCCCAAGACCTGTGTTTAACTTTGTGTGAAAATCTTTCCGCTGCAAC 471
DB |||
OY 421 ACACAACTCCAAGAAACCTCAACAGGTACTGTTCGCTGCAAGGTGAAGCAGACTTG 480
DB |||
OY 472 ACACAACTCCAAGAAACCTCAACAGGTACTGTTCGCTGCAAGGTGAAGCAGACTTG 531
DB |||
OY 481 GCTCAAGGAGGTGACGTGTAAATGACACCAAGGAGCTCTCTTGAATTTCAACATTTTG 540
DB |||
OY 532 GCTCAAGGAGGTGACGTGTAAATGACACCAAGGAGCTCTCTTGAATTTCAACATTTTG 591
DB |||
OY 541 CAATCTGAATTTGGGATTTTGAAGCTTTGGAAGAGAGACACTTACCAAAAAATAAA 600
DB |||
OY 592 CAATCTGAATTTGGGATTTTGAAGCTTTGGAAGAGAGACACTTACCAAAAAATAAA 651
DB |||
```

```
OY 601 TACATACCTCAGCAAGACGACTAGAGACAAATCGTGAATTTTACCATAACCAATT 660
DB |||
OY 652 TACATACCTCAGCAAGACGACTAGAGACAAATCGTGAATTTTACCATAACCAATT 711
DB |||
OY 661 GGAACAAACACAGCAAGATTCAGATTTTCAAGCTCTTAGAGATTGCCGTGCTAGAGATG 720
DB |||
OY 712 GGAACAAACACAGCAAGATTCAGATTTTCAAGCTCTTAGAGATTGCCGTGCTAGAGATG 771
DB |||
OY 721 TATGAAATCCGGTTGACCCCGGCAAGGACAGGGAAGGACAGATCAATCTGGCCGTT 780
DB |||
OY 772 TATGAAATCCGGTTGACCCCGGCAAGGACAGGGAAGGACAGATCAATCTGGCCGTT 831
DB |||
OY 781 GCCAACACGGGAATTTAGTGTTCACAGGTTTCACTAAGATCAATGCTTCAACTGGGCC 840
DB |||
OY 832 GCCAACACGGGAATTTAGTGTTCACAGGTTTCACTAAGATCAATGCTTCAACTGGGCC 891
DB |||
OY 841 AAGTGGCGGAGCTGAGCTTCAAGAGGAAGCCCTTCTCATCAAGCTCCGGCCAGATGCC 900
DB |||
OY 892 AAGTGGCGGAGCTGAGCTTCAAGAGGAAGCCCTTCTCATCAAGCTCCGGCCAGATGCC 951
DB |||
OY 901 AATAGGCGGTACAGGATACCTTGAATTCCTGATGGCCAGTCCGGATTTCTGCAAGTCC 960
DB |||
OY 1021 AAGCCCAAGCCGCTCTCTTTAGCCGGGGGTATCATTTCCGTTCAATGCTGGACTCAG 1080
DB |||
OY 1072 AAGCCCAAGCCGCTCTCTTTAGCCGGGGGTATCATTTCCGTTCAATGCTGGACTCAG 1131
DB |||
OY 1081 AAGCAGGTTCTGACATGTTTAAAGAGAGACATTAAGAAAGTGCAGTTGAAAGAAAG 1140
DB |||
OY 1132 AAGCAGGTTCTGACATGTTTAAAGAGAGACATTAAGAAAGTGCAGTTGAAAGAAAG 1191
DB |||
OY 1141 CACAGCAAGATTCATTTATCCGAGCCCTTGCTTCAACAGCCTTACAGAACTGAATTCGAA 1200
DB |||
OY 1192 CACAGCAAGATTCATTTATCCGAGCCCTTGCTTCAACAGCCTTACAGAACTGAATTCGAA 1251
DB |||
OY 1201 GTGCTGAGAGACTCTCAGCAGAGCACAGCCTTACATTTGGAGAAAGTGGCCGAATTCGA 1260
DB |||
OY 1252 GTACTGAGAGACTCTCAGCAGAGCACAGCCTTACATTTGGAGAAAGTGGCCGAATTCGA 1311
DB |||
OY 1261 GGGGGCCAGAGCTGCGCGGAGAAAGAAACCGAAGGTTTCCGCGGGGAGCCGGGGTGG 1320
DB |||
OY 1312 GGGGGCCAGAGCTGCGCGGAGAAAGAAACCGAAGGTTTCCGCGGGGAGCCGGGGTGG 1371
DB |||
OY 1321 CACCGAGCCCTGCGCGGAGAAAGAAACCGGAGTTAACAGACAGCGGAGCGAGCCGCC 1380
DB |||
OY 1372 CACCGAGCCCTGCGCGGAGAAAGAAACCGGAGTTAACAGACAGCGGAGCGAGCCGCC 1431
DB |||
OY 1381 TGGGGGCCCAAGAGAGAGAGAGAGGTCTTTAAGATNAGAACCAAGAGATTAACCT 1440
DB |||
OY 1432 TGGGGGCCCAAGAGAGAGAGAGAGGTCTTTAAGATNAGAACCAAGAGATTAACCT 1491
DB |||
OY 1441 CAGCCCGGCAAGCAAGCAAGGCTCCCTGACTGGCAAGTCACTTCTTCCAGAGCTGTCT 1500
DB |||
OY 1492 CAGCCCGGCAAGCAAGCAAGGCTCCCTGACTGGCAAGTCACTTCTTCCAGAGCTGTCT 1551
DB |||
OY 1501 GTGAATCTGCAAGGAGGAGTGGCCCTTGCAAGCTGACTTGTCTTCCAACTGAGCCCC 1560
DB |||
OY 1552 GTGAATCTGCAAGGAGGAGTGGCCCTTGCAAGCTGACTTGTCTTCCAACTGAGCCCC 1611
DB |||
OY 1561 GACACCAAGCAGGCTCTTCTCTTGAATCAAGCCGCTGTGAAATGACAGGCTTGGCCCCGG 1620
DB |||
OY 1612 GACACCAAGCAGGCTCTTCTCTTGAATCAAGCCGCTGTGAAATGACAGGCTTGGCCCCGG 1671
DB |||
OY 1621 ACAGAGATGAGAGTGGGGCCGAGAGAAAGATTTCCCAACTGATTAACGCTACTATTA 1680
DB |||
OY 1672 ACAGAGATGAGAGTGGGGCCGAGAGAAAGATTTCCCAACTGATTAACGCTACTATTA 1731
DB |||
OY 1681 GCTAAGGAAGTGTCTAACACCGAGGAAATATCTGAAGGATCTCGAAGTTACTCTTGG 1740
DB |||
```

Db 1732 GCTAAGAGAGTGTCTACCAACGAGCGAACTATCTGAAGATCTCGAAGTATCACTTGG 1791
 Qy 1741 TGGTTTCAGAGCACTAGTAGCAAAAGAGAGCCCATGCGGAGCACTGAAAGTCTCATTA 1800
 Db 1792 TGGTTTCAGAGCGCACTAGAGCAAAAGAGAGCGCATGCGAAGACACTGAAAGTCTCATTA 1851
 Qy 1801 TTCGCAATTTTGAACCTTTGCACAATTTTCACTAATTTTCTCAAGGAAATTTAGCAA 1860
 Db 1852 TTCCCAATTTTGAACCTTTGCACAATTTTCACTAATTTTCTCAAGGAAATTTAGCAA 1911
 Qy 1861 CGACTTGCCTGTGGAGAGCGCTCAATATGCCAATTCAGAGATTAACCAAGAAATCGGC 1920
 Db 1912 CGACTTGCCTGTGGAGAGCGCTCAATATGCCAATTCAGAGATTAACCAAGAAATCGGC 1971
 Qy 1921 GATGTCATGTGTAAGAACTTTCAGGGCATAGACACTGCGGCTCACTGTGGAAGCAC 1980
 Db 1972 GATGTCATGTGTAAGAACTTTCAGGGCATAGACACTGCGGCTCACTGTGGAAGCAC 2031
 Qy 1981 AGCGAGGCGCTTGGAGGCGCTGAGAAATGGAATGCAAGAGCTCCGGGGGCGTGGAGAACTTC 2040
 Db 2032 AGCGAGGCGCTTGGAGGCGCTGAGAAATGGAATGCAAGAGCTCCGGGGGCGTGGAGAACTTC 2091
 Qy 2041 TGCAGAGACTTTGAGCTGCAAGAAAGTGTGTTACCTACCGCTCAACAACCTTCTCTGCGG 2100
 Db 2092 TGCAGAGACTTTGAGCTGCAAGAAAGTGTGTTACCTACCGCTCAACAACCTTCTCTGCGG 2151
 Qy 2101 CCACTGCAACCGGCTCATGCACTACCAAGAGAGTCTTGAAGCGGCTGTGCAAAACCAACCGG 2160
 Db 2152 CCACTGCAACCGGCTCATGCACTACCAAGAGAGTCTTGAAGCGGCTGTGCAAAACCAACCGG 2211
 Qy 2161 CGAGGCAAGCGGCTTCAAGGGAATGCGAGCGGCTTGGCAGAGATTCACGAGATGATG 2220
 Db 2212 CGAGGCAAGCGGCTTCAAGGGAATGCGAGCGGCTTGGCAGAGATTCACGAGATGATG 2271
 Qy 2221 GCACAGCTCCAGCGGTACGATGATCAAGATGAGAAATTTCCAGAGCTGCAAGAACTCAAG 2280
 Db 2272 GCACAGCTCCAGCGGTACGATGATCAAGATGAGAAATTTCCAGAGCTGCAAGAACTCAAG 2331
 Qy 2281 AAAAGATTGATTGGCATTTGCACTTTGTGTGTTCCGGGAAAGGAGTTTCACTCCGTGCGG 2340
 Db 2332 AAAAGATTGATTGGCATTTGCACTTTGTGTGTTCCGGGAAAGGAGTTTCACTCCGTGCGG 2391
 Qy 2341 AGCTCAGCAAGCTCCGGGGAAGGGGCTCCAGCAGGCAATGTTCTTCCGTTCACAGAC 2400
 Db 2392 AGCTCAGCAAGCTCCGGGGAAGGGGCTCCAGCAGGCAATGTTCTTCCGTTCACAGAC 2451
 Qy 2401 GTCTGCTATATACAGAGCGGCGGCTGACGGCTCCAAATCAGTTTAAAGTCCACGGGCAAG 2460
 Db 2452 GTCTGCTATATACAGAGCGGCGGCTGACGGCTCCAAATCAGTTTAAAGTCCACGGGCAAG 2511
 Qy 2461 GTCTGCTATATAGCATGACGATTTGAGAGAGCGAAGACGAGTGGGCGGCTGCCCACTGC 2520
 Db 2512 GTCTGCTATACGCGCATGACGATTTAAGAGAGCGAAGACGAGTGGGCGGCTGCCCACTGC 2571
 Qy 2521 CTGACCCCTCCGGGGGCGAGGCGCATGATCATGATGCGGCGGCTTCTCGGTCTGAGATG 2580
 Db 2572 CTGACCCCTCCGGGGGCGAGGCGCATGATCATGATGCGGCGGCTTCTCGGTCTGAGATG 2631
 Qy 2581 GAGAAAGTGGGTTGAGAGCAATCCAGATGCGCATTTGACCTGCGGAGAGAGACAGACGCC 2640
 Db 2632 GAGAAAGTGGGTTGAGAGCAATCCAGATGCGCATTTGACCTGCGGAGAGAGACAGACGCC 2691
 Qy 2641 GCCCCTGAGTCTTGGCGCAGAGCGCCCTGACCAACAAGTCCCTTGAATGAGACACCGG 2700
 Db 2692 GCCCCTGAGTCTTGGCGCAGAGCGCCCTGACCAACAAGTCCCTTGAATGAGACACCGG 2751
 Qy 2701 GCTGACCAAGAGTCAAGAGTACCTGAGCGCTCGGAGCAATCGGTGAGAGCGGCAAGGCG 2760
 Db 2752 GCTGACCAAGAGTCAAGAGTACCTGAGCGCTCGGAGCAATCGGTGAGAGCGGCAAGGCG 2811
 Qy 2761 CCGCACCGGCGCAACAATGATGACGCTGTGTGCGACCGCAACAACGAGCGTCTCCATG 2820

Db 2812 CCGCACCGGCGCAACAATGATGTCATGTGTGCTGGACCGCAACACAGCGTCTCCATG 2871
 Qy 2821 GTGGACTTTCAGCATGCCATGTGGAATTCATGTTTCTGGAAAACCTGCTGAGGAAATTCCAA 2880
 Db 2872 GTAGACTTTCAGCGCTTTCAGATGAGAAATGATGTTGTGGGAACCTGCTGAGGAAATTCCAA 2931
 Qy 2881 AACAGCAAGGGGAGGAGAAAGCTGTGGTGTGTTTCAAAACTTCTGCTGTTCTTAC 2940
 Db 2932 AACAGCAAGGGGAGGAGAAAGCTGTGGTGTGTTTCAAAACTTCTGCTGTTCTTAC 2991
 Qy 2941 AAATCAACACGAGACATATCCCTTGTGCCAGCTGCTGTGCTGAGTATCCTGCTACC 3000
 Db 2992 AAGTCAACACGAGACATATCCCTGCGCAGCGCTGCTGTGCTGAGTATCCTGCTACC 3051
 Qy 3001 ATCCCTCTGAGTCCGAGAAATCCAGAAAGATACGTGTTCAAGTTCAGCTTCAAGTCC 3060
 Db 3052 ATCCCTCTGAGTCCGAGAAATCCAGAAAGATACGTGTTCAAGTTCAGCTTCAAGTCC 3111
 Qy 3061 CACGTCTACTACTTTCAGGGGCGGAAAGAGTACAGCTTCCGAAAGTGTGATGAAATGATC 3120
 Db 3112 CACGTCTACTACTTTCAGGGGCGGAAAGAGTACAGCTTCCGAAAGTGTGATGAAATGATC 3171
 Qy 3121 CGCAGTCCACCAAGCTCTGCTGCGACCCCAAGTGTGAGCCCAAGAGTCTTGTG 3180
 Db 3172 CGCAGTCCACCAAGCTCTGCTGCGAGTCCAGATGTCAGATGTCAGCAAAAGTCTCTGATG 3231
 Qy 3181 TATTGATGCGCGGACACATCTGTTTCCGATGAGTGTGCTTTCTGGAAGACGTTTCTTT 3240
 Db 3232 TATTGATGCGCGGACACATCTGTTTCCGATGAGTGTGCTTTCTGGAAGACGTTTCTTT 3291
 Qy 3241 CTTCTGATTTAATGAAGCTGTTGAAATTAACACCTGTCTGAAATCAAAAATCATGAGCTT 3300
 Db 3292 CTTCTGATTTAATGAAGCTGTTGAAATTAACACCTGTCTGAAATCAAAAATCATGAGCTT 3351
 Qy 3301 CCCAGAGCTCTCTGCTGTCACAGCGCGG-TTTTTAAACCCGACCTCTCAGCGTTTGA 3359
 Db 3352 CCCAGAGCTCTCTGCTGTCACAGCGCGGTTTAAACCCGACCTCTCAGCGCTGGA 3411
 Qy 3360 ATGAAACAGGCTCCCACTCCAGTCTGTGATCCGCTGGGGGCGCTTTCTTAAGCTAGT 3419
 Db 3412 ACMAACAGGCTCCCACTCCAGTCTGTGATCCGCTGGGGGCGCTTTCTTAAGCTAGT 3471
 Qy 3420 GCCAGATTAAACATTTGTCATT 3442
 Db 3472 GCCAGATTAAACATTTGTCATT 3494

RESULT 2
 AY406167 3138 bp DNA linear GSS 15-DEC-2003
 LOCUS
 DEFINITION Homo sapiens PARP1 gene, VIRUTUAL TRANSCRIPT, partial sequence.
 ACCESSION AY406167
 VERSION AY406167.1 GI:39762141
 KEYWORDS
 SOURCE GSS.
 ORGANISM Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Hominoidea; Homo.
 1 (bases 1 to 3138)
 REFERENCE
 AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejaritwal,A.,
 Todd,M.A., Tanenbaum,D.M., Civeille,D.R., Lu,F., Murphy,B.,
 Ferriere,S., Wang,G., Zheng,X.H., White,T.J., Shtinsky,J.J.,
 Adams,M.D. and Cargill,M.
 TITLE
 Inferring nonneutral evolution from human-chimp-mouse orthologous
 gene trios
 JOURNAL Science 302 (5652), 1960-1963 (2003)
 PUBMED 14671302
 REFERENCE 2 (bases 1 to 3138)
 AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejaritwal,A.,
 Todd,M.A., Tanenbaum,D.M., Civeille,D.R., Lu,F., Murphy,B.,
 Ferriere,S., Wang,G., Zheng,X.H., White,T.J., Shtinsky,J.J.,

Adams,M.D. and Cargill,M.
 Direct Submission
 Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
 Rockville, MD 20850, USA
 This sequence was made by sequencing genomic exons and ordering
 them based on alignment.
 Location/Qualifiers
 1. 3138
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 <1..>3138
 /gene="FARP1"
 /locus_tag="HCM2467"

ORIGIN

Query Match	91.1%;	Score 3136.4;	DB 14;	Length 3138;
Best Local Similarity	99.9%;	Pred. No. 0;		
Matches 3137;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;
QY	49	ATGGGAGAAATGAGCAGAGGCGGACCCAGATCACTGGGGGCCCCGGAAATTGG	108	
DB	1	ATGGGAGAAATGAGCAGAGGCGGACCCAGATCACTGGGGGCCCCGGAAATTGG	60	
QY	109	GGGATCAGTACCTTGGAGCGTGACAGAACCCGCCCAACACTTCAGAGAAACTCGTG	168	
DB	61	GGGATCAGTACCTTGGAGCGTGACAGAACCCGCCCAACACTTCAGAGAAACTCGTG	120	
QY	169	TCCATCAAAATCCAGATGCTGATGACACCCAGAGAGCATTTGAAGTTCCACAAAGACT	228	
DB	121	TCCATCAAAATCCAGATGCTGATGACACCCAGAGAGCATTTGAAGTTCCACAAAGACT	180	
QY	229	CCGGGAAGGTGCTGTGATGATGATTTTCAACACCTCACTCGTGGAGGTGACTAT	288	
DB	181	CCGGGAAGGTGCTGTGATGATGATTTTCAACACCTCACTCGTGGAGGTGACTAT	240	
QY	289	TTTGCCCTGAGTTTCTGATCACAAAAAGATCACGGTGTGGTGTGATCTCTTAAACC	348	
DB	241	TTTGCCCTGAGTTTCTGATCACAAAAAGATCACGGTGTGGTGTGATCTCTTAAACC	300	
QY	349	ATTGTGAAAACAGTTTGAAAGGCCAAAGCACTTTGTTGTTAAGTTTGTGTAATCTTT	408	
DB	301	ATTGTGAAAACAGTTTGAAAGGCCAAAGCACTTTGTTGTTAAGTTTGTGTAATCTTT	360	
QY	409	CCGCCCTGACACACAACTCCAGAAAGACTCACAAGGATCCTGTTCGGCGTGCAGTG	468	
DB	361	CCGCCCTGACACACAACTCCAGAAAGACTCACAAGGATCCTGTTCGGCGTGCAGTG	420	
QY	469	AAGCAGGACTTGGCTCAAGCAGGTGACGTGTATGACACACGCGAGCTCTTGATT	528	
DB	421	AAGCAGGACTTGGCTCAAGCAGGTGACGTGTATGACACACGCGAGCTCTTGATT	480	
QY	529	TCAACATTTGTGCAATCTGAGATTGGGGATTTTGAAGCCTTGGACAGAGCACTTA	588	
DB	481	TCAACATTTGTGCAATCTGAGATTGGGGATTTTGAAGCCTTGGACAGAGCACTTA	540	
QY	589	GCAAAAATAAATATACATCTCAGCAAGACGCACTAGAGACAAATCGTGAATTTAC	648	
DB	541	GCAAAAATAAATATACATCTCAGCAAGACGCACTAGAGACAAATCGTGAATTTAC	600	
QY	649	CATTAACCATTTGACAAACACCAAGCAGATCAGATTTTCAGGCTCTAGAGATTGCCGT	708	
DB	601	CATTAACCATTTGACAAACACCAAGCAGATCAGATTTTCAGGCTCTAGAGATTGCCGT	660	
QY	709	CGGCTAGAGATGTATGGAATTCGGGTGGACCCGGCCAAAGACAGGGAAGGCAGGAATC	768	
DB	661	CGGCTAGAGATGTATGGAATTCGGGTGGACCCGGCCAAAGACAGGGAAGGCAGGAATC	720	
QY	769	AATCTGGCCGTGGCAACGAGGAATTTAGTGTTCAGGGTTTCACTAAGATCAATGCC	828	
DB	721	AATCTGGCCGTGGCAACGAGGAATTTAGTGTTCAGGGTTTCACTAAGATCAATGCC	780	
QY	829	TTCAACTGGGCCAAGGTGGGAAAGCTGAGCTTCAGAGGAAGGCTTTCTCAATCAAGCTC	888	

DB	781	TTCAACTGGGCCAAGGTGGGAAAGCTGAGCTTCAGAGGAACGCTTCTCAATCAAGCTC	840	
QY	889	CGGCCAGATGCCAATATGTCCTTACAGAGTACTCTTGAAATTCCTATAGGCCAGTCGGGAT	948	
DB	841	CGGCCAGATGCCAATATGTCCTTACAGAGTACTCTTGAAATTCCTATAGGCCAGTCGGGAT	900	
QY	949	TTCTGCAAGTCTTCTGGAATACTGTGTGAACATCATGCTCTTTAGACTTTTGA	1008	
DB	901	TTCTGCAAGTCTTCTGGAATACTGTGTGAACATCATGCTCTTTAGACTTTTGA	960	
QY	1009	GAGCCCAACCAAGCCCAAGCCGCTCTTTAGCCGGGGTCAATATTGCTTCACT	1068	
DB	961	GAGCCCAACCAAGCCCAAGCCGCTCTTTAGCCGGGGTCAATATTGCTTCACT	1020	
QY	1069	GGTGGACTCAAGAACAGTTTCTGACTATGTATTAAGAAAGAGACATTAAGAAAGTCAG	1128	
DB	1021	GGTGGACTCAAGAACAGTTTCTGACTATGTATTAAGAAAGAGACATTAAGAAAGTCAG	1080	
QY	1129	TTTGAAGGAAGCAGCAGCAAGATTCAATTCATCCGAGGCTTGTTCAAGCTTACAGAA	1188	
DB	1081	TTTGAAGGAAGCAGCAGCAAGATTCAATTCATCCGAGGCTTGTTCAAGCTTACAGAA	1140	
QY	1189	CTGAATTGGAAGTCTGAGAGCATCTCAGACAGACCAAGCTTACATTTGAGAAAGT	1248	
DB	1141	CTGAATTGGAAGTCTGAGAGCATCTCAGACAGACCAAGCTTACATTTGAGAAAGT	1200	
QY	1249	GCCGAATCTCCAGGGGGCCAGAGCTGCGGCGAGGAAGAAAGCAAGAGTTTCCCGGGG	1308	
DB	1201	GCCGAATCTCCAGGGGGCCAGAGCTGCGGCGAGGAAGAAAGCAAGAGTTTCCCGGGG	1260	
QY	1309	GAGCCGGGTGCGACCCGAGCCCTGCGCGAGGAAGCCCGGGGTAAACAAGACAGCG	1368	
DB	1261	GAGCCGGGTGCGACCCGAGCCCTGCGCGAGGAAGCCCGGGGTAAACAAGACAGCG	1320	
QY	1369	GAGGAGCCGCTCGGCGCCACGAGGAAGAGAGAGGTGTTAAGATTGACCCAG	1428	
DB	1321	GAGGAGCCGCTCGGCGCCACGAGGAAGAGAGAGGTGTTAAGATTGACCCAG	1380	
QY	1429	GAGATTAACCTCAAGCCCGGAGCAAGACAGGCTCCCTGATCGGAGCTCAACCTT	1488	
DB	1381	GAGATTAACCTCAAGCCCGGAGCAAGACAGGCTCCCTGATCGGAGCTCAACCTT	1440	
QY	1489	TCCGAGCTGTCTGTAACTCGCAGGGGGAGTGGCCCTGCAAGCTGACTTGTCTCC	1548	
DB	1441	TCCGAGCTGTCTGTAACTCGCAGGGGGAGTGGCCCTGCAAGCTGACTTGTCTCC	1500	
QY	1549	AACTGAAGCCCGACACACAGAGGCTCTCTCTTGATAGCCCGCTGTAATGACAG	1608	
DB	1501	AACTGAAGCCCGACACACAGAGGCTCTCTCTTGATAGCCCGCTGTAATGACAG	1560	
QY	1609	GCTGCCCCCGGAGGAGTAGAGATGAGGGCGGAGGAAGAGATTCCCACTGATAA	1668	
DB	1561	GCTGCCCCCGGAGGAGTAGAGATGAGGGCGGAGGAAGAGATTCCCACTGATAA	1620	
QY	1669	GCGTACTTATAGTAAGGAAGTGTCTACACCGAGCGCAATATCTGAAGATCTGAA	1728	
DB	1621	GCGTACTTATAGTAAGGAAGTGTCTACACCGAGCGCAATATCTGAAGATCTGAA	1680	
QY	1729	GTTATCATCTTGTGTTTCAAGCAAGTAGCAAAAGAGACGCAATGCGGAGACATG	1788	
DB	1681	GTTATCATCTTGTGTTTCAAGCAAGTAGCAAAAGAGACGCAATGCGGAGACATG	1740	
QY	1789	AAAAGTCTATTTCCCGAATTTTGAACCTTGGCAAAATTCATTAATTTTCTCAAG	1848	
DB	1741	AAAAGTCTATTTCCCGAATTTTGAACCTTGGCAAAATTCATTAATTTTCTCAAG	1800	
QY	1849	GAAATTGAGCAACGACTTGCCCTGTGGGAAGGCCCTCAATAGCCCAATCAGAGATTAC	1908	
DB	1801	GAAATTGAGCAACGACTTGCCCTGTGGGAAGGCCCTCAATAGCCCAATCAGAGATTAC	1860	
QY	1909	CAGAATGCGGAGTGTACTGCTGAAGAACATTCAGGGCATGAAGCACTTGGCGCTCAC	1968	

```

Db      1861 CAAGAATCGCGATGTCATGTAAGAACATTCAAGGCGATGAAGCACTGCGCGCTCAC 1920
Qy      1969 CTGTGGAAGACAGCGAGGCTTGGAGAGCCCTTGGAGAAATGGAATCAAGAGCTCCCGCGG 2028
Db      1921 CTGTGGAAGACAGCGAGGCTTGGAGAGCCCTTGGAGAAATGGAATCAAGAGCTCCCGCGG 1980
Qy      2029 CTGGAAGAACTTTCAGAGAGACTTGGAGCTCAGAGAGGTGTGTTACCTACCGCTCAACACC 2088
Db      1981 CTGGAAGAACTTTCAGAGAGACTTGGAGCTCAGAGAGGTGTGTTACCTACCGCTCAACACC 2040
Qy      2089 TTCCTCTCTGAGGCACTGCAACCGGCTTCACTCAAGAGAGTCTTGAAGCGGCTGTGC 2148
Db      2041 TTCCTCTCTGAGGCACTGCAACCGGCTTCACTCAAGAGAGTCTTGAAGCGGCTGTGC 2100
Qy      2149 AACACACACCCCGGAGCCAGCCGACTTCAAGGAGCTGCGAGCCGCTTTGGCAGAGATC 2208
Db      2101 AACACACACCCCGGAGCCAGCCGACTTCAAGGAGCTGCGAGCCGCTTTGGCAGAGATC 2160
Qy      2209 ACGGAGATGGTGGCAGAGCTCCAGCGGTACGATGATCAAGATGGAATTTCCAGAAAGCTG 2268
Db      2161 ACGGAGATGGTGGCAGAGCTCCAGCGGTACGATGATCAAGATGGAATTTCCAGAAAGCTG 2220
Qy      2269 CACGAACTCAAGAAAGATTGATTGCGATTGACAACTTGTGTTCCGGAGAGGAGTTTC 2328
Db      2221 CACGAACTCAAGAAAGATTGATTGCGATTGACAACTTGTGTTCCGGAGAGGAGTTTC 2280
Qy      2329 ATCCGCTCTGGGAGCTTCAAGAGCTTCTGGGGGAAAGGGGCTCCAGCAGCGCATGTTCTTC 2388
Db      2281 ATCCGCTCTGGGAGCTTCAAGAGCTTCTGGGGGAAAGGGGCTCCAGCAGCGCATGTTCTTC 2340
Qy      2389 CTGTTTCAAGACGCTCTGTATATACAGAGCGGGGGGTGACGGGCTTCAATCAGTTTAA 2448
Db      2341 CTGTTTCAAGACGCTCTGTATATACAGAGCGGGGGGTGACGGGCTTCAATCAGTTTAA 2400
Qy      2449 GTTCACGGGAGCTCCCGCTCTATATGCGATGACGATTGAGAGAGCGAAGACGAGTGGGG 2508
Db      2401 GTTCACGGGAGCTCCCGCTCTATATGCGATGACGATTGAGAGAGCGAAGACGAGTGGGG 2460
Qy      2509 GTGCCCCACCTGCTGACCTTCCGGGGGCGAGCGGACGTCCATCATCTGTGGCCGCAATTC 2568
Db      2461 GTGCCCCACCTGCTGACCTTCCGGGGGCGAGCGGACGTCCATCATCTGTGGCCGCAATTC 2520
Qy      2569 CGGTCGAGATGAGAGAGTGGTTGAGAGATCAAGATGAGCAATTCGTCGCGAGAAAG 2628
Db      2521 CGGTCGAGATGAGAGAGTGGTTGAGAGATCAAGATGAGCAATTCGTCGCGAGAAAG 2580
Qy      2629 AGCAGCAGCCCGCCCTGAGTTCCTGCGCAGAGCCCGCTTGAACAACAAGTCCCTGAT 2688
Db      2581 AGCAGCAGCCCGCCCTGAGTTCCTGCGCAGAGCCCGCTTGAACAACAAGTCCCTGAT 2640
Qy      2689 GAAGCCACCGCGGCTGACCAAGAGTCAAGAGATGACTGAGCGCTTCGCGCAGATGCTG 2748
Db      2641 GAAGCCACCGCGGCTGACCAAGAGTCAAGAGATGACTGAGCGCTTCGCGCAGATGCTG 2700
Qy      2749 GAGCGCAGGCGCCCGCAGCGCGGCAACAATGATGACGTGTGCTGGCAGCGCAACCC 2808
Db      2701 GAGCGCAGGCGCCCGCAGCGCGGCAACAATGATGACGTGTGCTGGCAGCGCAACCC 2760
Qy      2809 AGCGTCTCAATGATGATTCAGCATTCGAGTGAAGATCAAGTGTCTGAAAACCTGCTG 2868
Db      2761 AGCGTCTCAATGATGATTCAGCATTCGAGTGAAGATCAAGTGTCTGAAAACCTGCTG 2820
Qy      2869 AGGAATTTCAAAAACAGCAACCGGTGGCAGAACTGTGGGTGTTCACAAACTTCTGC 2928
Db      2821 AGGAATTTCAAAAACAGCAACCGGTGGCAGAACTGTGGGTGTTCACAAACTTCTGC 2880
Qy      2929 CTGTTCTTACAAATACACACAGAGCAATCATCCCTTGCAGAGCTGCTGTCTGCGC 2988
Db      2881 CTGTTCTTACAAATACACACAGAGCAATCATCCCTTGCAGAGCTGCTGTCTGCGC 2940
Qy      2989 TACTGCTCACATCCCTCTGATGTCGAGAACATCCAGAAAGACTACGTGTCAAGCTG 3048
Db      2941 TACTGCTCACATCCCTCTGATGTCGAGAACATCCAGAAAGACTACGTGTCAAGCTG 3000

```

```

Qy      3049 CACTTCAAGTCCCAAGTCTACTTCACTTCAAGGCGGAAAGGAGTACAGCTTCCAAAAGTGG 3108
Db      3001 CACTTCAAGTCCCAAGTCTACTTCACTTCAAGGCGGAAAGGAGTACAGCTTCCAAAAGTGG 3060
Qy      3109 ATGGAAGTATCCGACAGTCCACACAGCTCTGCTCGCGACCCCACTGTGTGAGCCACAA 3168
Db      3061 ATGGAAGTATCCGACAGTCCACACAGCTCTGCTCGCGACCCCACTGTGTGAGTACAA 3120
Qy      3169 GAGTCTCTTGTGTATGGA 3186
Db      3121 GAGTCTCTTGTGTATGGA 3138

RESULT 3
AY406168      3138 bp      DNA      linear      GSS 15-DEC-2003
LOCUS      Pan troglodytes FARP1 gene, VIRUTAL TRANSCRIPT, partial sequence,
DEFINITION      genomic survey sequence.
ACCESSION      AY406168
VERSION      AY406168.1 GI:39762142
KEYWORDS      GSS.
SOURCE      Pan troglodytes (chimpanzee)
ORGANISM      Pan troglodytes
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Pan.
REFERENCE      1 (bases 1 to 3138)
AUTHORS      Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejaritwal,A.,
      Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
      Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Shtinsky,J.U.,
      Adams,M.D. and Cargill,M.
      Inferring nonneutral evolution from human-chimp-mouse orthologous
      gene clusters
      Science 302 (5652), 1960-1963 (2003)
JOURNAL      PubMed
PUBMED      14671302
AUTHORS      Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejaritwal,A.,
      Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
      Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Shtinsky,J.U.,
      Adams,M.D. and Cargill,M.
      Direct Submission
      Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
      Rockville, MD 20850, USA
      This sequence was made by sequencing genomic exons and ordering
      them based on alignment.
FEATURES      Location/Qualifiers
      source      1..3138
                  /organism="Pan troglodytes"
                  /mol_type="genomic DNA"
                  /db_xref="taxon:9598"
                  <1..>3138
                  /gene="FARP1"
                  /locus_tag="HCM2467"
ORIGIN
Query Match      86.2%; Score 2968.2; DB 14; Length 3138;
Best Local Similarity 94.8%; Pred. No. 0;
Matches 2976; Conservative 0; Mismatches 162; Indels 0; Gaps 0;
Qy      49 ATGGAGAAATAGAGCAGAGCGGACCCCAAGATACAGACTGGGGCCCCGGAAAAATTGG 108
Db      1 ATGGAGAAATAGAGCAGAGCGGACCCCAAGATACAGACTGGGGCCCCGGAAAAATTGG 60
Qy      109 GGGATCAGTACTTGGAACTGGACAGAGCCGCCCAACACTTTCAGAAAACTCGTG 168
Db      61 GGGATCAGTACTTGGAACTGGACAGAGCCGCCCAACACTTTCAGAAAACTCGTG 120
Qy      169 TGCATAAATATCAGATGCTGATGACACCCAGAGAGGCAATTGAGATTCCAAAGAGCT 228
Db      121 TGCATAAATATCAGATGCTGATGACACCCAGAGAGGCAATTGAGATTCCANNNNNNNN 180
Qy      229 CCTGGGAAGTCTGCTGATCAAGTTTCAACCACTCAACCTCGTGGAGGTGACTAT 288

```

[illegible]

Dd	1261	GAGCCGGGGGTGCAACCCGAGCCCTGCGCCGAGGAGAAAGCCCGCGGGTAAACAGACAGGCG	1320
Qy	1369	GACGAGCCCGCTTCGCGGCCCAACGAGAGAAAGAGAGGTCTGTTAAGATAGAGCCAG	1428
Dd	1321	GACGAGACCGGCTCGCGGCCCAACGAGAGAAAGAGAGAGGTCTGTTAAGATAGAGCCAG	1380
Qy	1429	CAGAGTAAACCTCAGAGCCCGCCGAGCCAAAGACAGAGGTCCCTGACTGGCACTCCCACTT	1488
Dd	1381	CAGGTAAACCTCAGAGCCCGCCGAGCCAAAGACAGAGGTCCCTGACTGGCACTCCCACTT	1440
Qy	1489	TCCGAGCTGTCTGTGAACCTCGACAGGGGAGATGGCCCTTGCCAAACGAGCTTGTCTCC	1548
Dd	1441	TCCGAGCTGTCTGTGAACCTCGACAGGGGAGATGGCCCTTGCCAAACGAGCTTGTCTCC	1500
Qy	1549	AACTGAGCCCGGACACCAAGCAGGCTCTCCCTGATCAAGCCCGCTGTGAATGACAG	1608
Dd	1501	AACTGAGCCCGGACACCAAGCAGGCTCTCCCTGATCAAGCCCGCTGTGAATGACAG	1560
Qy	1609	GCTGCGCCCGGAGCAGACGATGAGAGTGAAGGGCCGAGAGAGATTCGCCACTGATAAA	1668
Dd	1561	GCTGCGCCCGGAGCAGACGATGAGAGTGAAGGGCCGAGAGAGATTCGCCACTGATAAA	1620
Qy	1669	GCGTACTTCATAGCTAAGAAAGTGTCTTACCAACGAGCGAAACATATCTGANGATCTCGAA	1728
Dd	1621	GCGTACTTCATAGCTAAGAGGNNNTCCNNNAACGAGGAAACATATCTGAAGGATCTTGAA	1680
Qy	1729	GTATATCACTTCGATGGTTTCAAGACACATGAGCAAAAGAGACGCAATGCGGAAAGCACTG	1788
Dd	1681	GTATATCACTTCGATGGTTTCAAGACACATGAGCAAAAGAGACGCAATGCGGAAAGCACTG	1740
Qy	1789	AAAAGTCTCATATTTCCGGAATTTTGAACCTTTGCACAAATTTCACTAAATTTTCTCAAG	1848
Dd	1741	AAAAGTCTCATATTTCCGGAATTTTGAACCTTTGCACAAATTTCACTAAATTTTCTCAAG	1800
Qy	1849	GAAATTGAGCAACGACTTGGCCCTGTGGGAGGCGCCCTCAATGCCAAATCCAGAGATTAC	1908
Dd	1801	GAAATTGAGCAACGACTTGGCCCTGTGGGAGGCGCCCTCAATGCCAAATCCAGAGATTAC	1860
Qy	1909	CAAAAGATCGGCGCATGTCTATGCTGAGAAACATTCAGGGCATGAGACACTTGGCGGCTCAC	1968
Dd	1861	CAAAAGATCGGCGCATGTCTATGCTGAGAAACATTCAGGGCATGAGACACTTGGCGGCTCAC	1920
Qy	1969	CTGTGGAAGCACAGCGGAGGCTTTGGAGGCCCTTGAGAAATGGAATCAAGACTCCCGGCGG	2028
Dd	1921	CTGTGGAAGCACAGCGGAGGCTTTGGAGGCCCTTGAGAAATGGAATCAAGACTCCCGGCGG	1980
Qy	2029	CTGAGAAACTTCTGACAGACTTTGAGCTCAGAGAGTGTGTTACTCAACGCTCAACACC	2088
Dd	1981	CTGAGAAACTTCTGACAGACTTTGAGCTCAGAGAGTGTGTTACTCAACGCTCAACACC	2040
Qy	2089	TTCTCTCTGCGGCCACTGCAACCGGCTCATGCACTAACAGAGTCTTGAGCGGCTGTGC	2148
Dd	2041	TTCTCTCTGCGGCCACTGCAACCGGCTCATGCACTAACAGAGTCTTGAGCGGNNGTGC	2100
Qy	2149	AAAACACACCGCGCGAGACCAAGCGCACTTCAAGGGACTGCCGAGCCGTTTGGACAGATC	2208
Dd	2101	AAAACACACCGCGCGAGNN	2160
Qy	2209	ACGAGAGATGATGAGCACAGCTCCACGATACGATGATCAAGATGAGAAATTTCCAGAACCTG	2268
Dd	2161	ACGAGAGATGATGAGCACAGCTCCACGATACGATGATCAAGATGAGAAATTTCCAGAACCTG	2220
Qy	2269	CACGAACTCAAGAAAGATTTGATTTGGATTTGATGACAAATCTTGTGGTTCGAGGAGGGAATTC	2328
Dd	2221	CACGAACTCAAGAAAGATTTGATTTGGATTTGATGACAAATCTTGTGGTTCGAGGAGGGAATTC	2280
Qy	2329	ATTCGCTGGGAGGCTCAGCAAGCTCTCGGAGGAGGAGGCTCAGCAGCCATGTTTCTC	2388
Dd	2281	ATTCGCTGGGAGGCTCAGCAAGCTCTCGGAGGAGGAGGCTCAGCAGCCATGTTTCTC	2340
Qy	2389	CTGTTCAACGACGTCTGTCTATAACGAGGCTCGGAGGCTGACGCGCTCAATCAAGTTTAA	2448
Dd	2341	CTGTTCAACGACGTCTGTCTATAACGAGGCTCGGAGGCTGACGCGCTCAATCAAGTTTAA	2400

QY 2449 GTCCAGGGGAGCTCCCGCTCTATGCGATGCAATTGAGGAGAGCGAAAGCGAGTGGGG 2508
DB 2401 GTCCAGGGGAGCTCCCGCTCTATGCGATGCAATTGAGGAGAGCGAAAGCGAGTGGGG 2460
QY 2509 GTGCCCCACCTGCTGACCTCCGGGGCGAGCGGAGTCACTCATCTGAGCGCGCAATTCT 2568
DB 2461 GTGCCCCACCTGCTGACCTCCGGGGCGAGCGGAGTCACTCATCTGAGCGCGCAATTCT 2520
QY 2569 CGGTCCGAGATGAGAAAGTGGGTGAGAGACATCCAGATGGCCATTGACTTGGGGAGAG 2628
DB 2521 CGGTCCGAGATGAGAAAGTGGGTGAGAGACATCCAGATGGCCATTGACTTGGGGAGAG 2580
QY 2629 AGCAGACGCCCGCCCTGAGTTCCTGCGCAGAGCGCCCTGAGCAACAGTCCCTGAT 2688
DB 2581 AGCAGACGCCCGCCCTGAGTTCCTGCGCAGAGCGCCCTGAGCAACAGTCCCTGAT 2640
QY 2689 GAAGCCACCGCGGCTGACAGAGTCAAGAGTACCTGAGCGCTTCGCCACATCGCTG 2748
DB 2641 GAAGCCACCGCGGCTGACAGAGTCAAGAGTACCTGAGCGCTTCGCCACATCGCTG 2700
QY 2749 GAAGCCACCGCGGCTGACAGAGTCAAGAGTACCTGAGCGCTTCGCCACATCGCTG 2808
DB 2701 GAAGCCACCGCGGCTGACAGAGTCAAGAGTACCTGAGCGCTTCGCCACATCGCTG 2760
QY 2809 AGGCTCCATGAGTGAATCAGCATGCGAGTGAAGTCAATGTTGCTGAAAACCTGCTG 2868
DB 2761 AGGCTCCATGAGTGAATCAGCATGCGAGTGAAGTCAATGTTGCTGAAAACCTGCTG 2820
QY 2869 AGGAAATTCAAAACAGCAACGCGTGGCAGAACTGTGGTGTGTCAAACTTCTGC 2928
DB 2821 AGGAAATTCAAAACAGCAACGCGTGGCAGAACTGTGGTGTGTCAAACTTCTGC 2880
QY 2929 CTGTGTTCTTCAAAATACACACAGACATCATCTCTTCCAGCTGCTCTGCTGGC 2988
DB 2881 CTGTGTTCTTCAAAATACACACAGACATCATCTCTTCCAGCTGCTCTGCTGGC 2940
QY 2989 TACTCGCTGACATCCCTCTGAGTCCGAGAACATCCGAAAGACTAGCTTCAAGCTG 3048
DB 2941 TACTCGCTGACATCCCTCTGAGTCCGAGAACATCCGAAAGACTAGCTTCAAGCTG 3000
QY 3049 CACTTCAGTCCGACGCTCTACTACTTCAAGGCGGAAAGCGAGTACAGTTCGAAAGTGG 3108
DB 3001 CACTTCAGTCCGACGCTCTACTACTTCAAGGCGGAAAGCGAGTACAGTTCGAAAGTGG 3060
QY 3109 ATGGAAGTATCCGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3168
DB 3061 ATGGAAGTATCCGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3120
QY 3169 GAGTCTCTGTTATTGA 3186
DB 3121 GAGTCTCTGTTATTGA 3138

RESULT 4
AY06169 3147 bp DNA linear GSS 15-DEC-2003
LOCUS AY06169
DEFINITION Mus musculus PARP1 gene, VIRUAL TRANSCRIPT, partial sequence.
ACCESSION AY06169
VERSION AY06169.1 GI:39762143
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Muridae; Murinae; Mus.
1 (bases 1 to 3147)
Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarival, A.,
Todd, M.A., Tanenbaum, D.M., Civeille, D.R., Lu, F., Murphy, B.,
Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Shtinsky, J.J.,
Adams, M.D. and Cargill, M.
Inferring nonneutral evolution from human-chimp-mouse orthologous

gene trios
Science 302 (5652), 1960-1963 (2003)
14671302
JOURNAL
PUBMED
REFERENCE
2 (bases 1 to 3147)
Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarival, A.,
Todd, M.A., Tanenbaum, D.M., Civeille, D.R., Lu, F., Murphy, B.,
Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Shtinsky, J.J.,
Adams, M.D. and Cargill, M.
TITL
JOURNAL
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT
This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
Location/Qualifiers
1..3147
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
<1..>3147
/gene="PARP1"
/locus_tag="Hm2467"
ORIGIN
Query Match 70.5%; Score 2427.4; DB 14; Length 3147;
Best Local Similarity 86.2%; Pred. No. 0;
Matches 2712; Conservative 0; Mismatches 426; Indels 9; Gaps 2;
QY 49 ATGGGAAATATGAGCAGAGCGCCAGCCAGATCCAGACTGGGGCCCCGGAAATTCG 108
DB 1 ATGGGAAATATGAGCAGAGCGCCAGCCAGATCCAGACTGGGGCCCCGGAAATTCG 60
QY 109 GGGATAGTACCTTGGAGAGTGGAGAGCGCGCCCAACCTTCAGAGAACTCGT 168
DB 61 GGGATAGTACCTTGGAGAGTGGAGAGCGCGCCCAACCTTCAGAGAACTTCG 120
QY 169 TCCATCAAAATCCAGATGCTGATGACACCCAGAGGCAATTTGAAGTTCCAGAAAGCT 228
DB 121 ACTGTCAAAATCCAGATGCTGATGACACCCAGAGGCAATTTGAAGTTCCAGAAAGCT 180
QY 229 CCGGGAAGTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 288
DB 181 CCGGGAAGTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
QY 289 TTGGGCTGAGTTCTGATCAAAAGATCAGGCTGAGTGGTGGATCTCTTAAACCC 348
DB 241 TTGGGCTGAGTTCTGATCAAAAGATCAGGCTGAGTGGTGGATCTCTTAAACCC 300
QY 349 ATTTGAAAACAGATTAGAGGCGAAAGCACTTTGTTAGTTGTTGTTGTTGTTGTTGTT 408
DB 301 ATTTGAAAACAGATTAGAGGCGAAAGCACTTTGTTAGTTGTTGTTGTTGTTGTTGTT 360
QY 409 CCGCTGACACACACATCCAGAGAACTCAAGATGATGATGATGATGATGATGATGATGAT 468
DB 361 CCGCTGACACACACATCCAGAGAACTCAAGATGATGATGATGATGATGATGATGATGAT 420
QY 469 AAGCAGACTTGGCTCAAGGAGTTGAGTGAATGACACAGGCAAGCTCTCTGATT 528
DB 421 AAGCAGACTTGGCTCAAGGAGTTGAGTGAATGACACAGGCAAGCTCTCTGATT 480
QY 529 TCACATTTGTCATCTGAGATTGGGATTTTGAATGAGCTTGGACAGAGCACTTA 588
DB 481 TCACATTTGTCATCTGAGATTGGGATTTTGAATGAGCTTGGACAGAGCACTTA 540
QY 589 GCAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 648
DB 541 GCAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 600
QY 649 CATTAACCATTTGAGCAACACACAGAGAAATGATGATGATGATGATGATGATGATGATGAT 708
DB 601 CATTAACCATTTGAGCAACACACAGAGAAATGATGATGATGATGATGATGATGATGATGAT 660
QY 709 CGGCTAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 768

Db 661 CGGCTAGAGATGATCGGAATCCGGTTACACCTCGCAAGGACGAGGAGGCAACAAGATC 720
Qy 769 AATTCGGCCGTTGGCCACAACGGGAATTTAGTGTTCAGGGTTTCACTAAGATCAATGCG 828
Db 721 AACTGGCTGTGGCCAACTGGAAATTTAGTGTTCAGAGCTTCAGGAAGATCAATGCT 780
Qy 829 TTCAACTGGGCGCAAGGTGGGAGAGCTAGACTTCAAGAGAAAGCGCTTTCATCAAGCTC 888
Db 781 TTTAACTGGGCGCAAGGTGGGAGAGCTAGACTTCAAGAGAAAGCGCTTTCATCAAGCTC 840
Qy 889 CGGCGAGATGCCAATATGTGCTACAGATACCTTGAAATTCCTGATGCGCAGTGGGAT 948
Db 841 AAGACCTGAGGTGATGTGCTATACAGGACACCTTGAGTTCCTGATGCGCGCCAGAC 900
Qy 949 TTCTGCAATCTCTTGGAAAAATCTGTGTGAACATCATGCGCTTCTTGAACCTTTTGA 1008
Db 901 TTCTGCAATCTTCTTGGAAAAATATGTGAGACACATGCTTCTTCAAGACTCTTCGAA 960
Qy 1009 GAGCCCAACCAAAAGCCCAAGCCCTCTCTTTAGCGGGGGTCAATCTTTCGGTTCACT 1068
Db 961 GAACCTAAGCCAAAGCCAAAGCTGTCTCTTCAAGCGAGGTGCTCTTCGGTTTACT 1020
Qy 1069 GGTGGACTCAAGAGAGGTTCTGACTATGTTAAAGAGAGACATAAGAGGTGAG 1128
Db 1021 GGTGGACTCAAGAGAGGTTCTGACTATGTTAAAGAGAGCGGCCCAAGAGAGGTGAG 1080
Qy 1129 TTTGAAAGAGACAGCAAGATTCATTTATCGAGGCTTGCTTCAAGGCTACAGAA 1188
Db 1081 TTTGAAAGAGACAGCAAGATTCATTCACAGAGGCTTGCTTCAAGGCTACAGAA 1140
Qy 1189 CTGAATTCGAAAGTGTGAGAGAGTCTCAGCAGAGACCAAGCTTACATTTGGAGAGGT 1248
Db 1141 CGGAATTCAGAGAGTGCAGAAACAGTCTCCACAGAGCGCAAGCTTACGTTGAGAGAGGT 1200
Qy 1249 GCGGAATCTTCAGAGGGGCCAGAGCTCCGGCCAGAGAAAGAACCAAGGTTTCGGCGGG 1308
Db 1201 ACCGAGTCTTCGGGTGGCCAGAGCTGCGCAGCAAGCCAGAGAAACAAAGCTGTACCTTG 1260
Qy 1309 GAGCGGGGGTGGACCCGAGGCGCTGCGCGAGAGAAAGCGCGGGGTAAACAAGAGGCG 1368
Db 1261 GAGTGGAGACTTCACCAAGGCGCGGCTTCCTTAAGGCTTTCAGGGAGCAAGGCTGCA 1320
Qy 1369 GACGAGCGGCTCGGCGGCCACGAGAGAA-----GAGAGAGAGGTCTTAAAGATAG 1422
Db 1321 GATGGAACCAAGCGGTGTGCGGCTGAGAAAGAGAGAGAGAGGGTGGCAAGAGCGGC 1380
Qy 1423 ACCCAGCAGATTAACCTCAGGCCCCCGCAGGCCCAAGCAGGCTCCTGTATGCGAGTCTT 1482
Db 1381 ATCCGGCCAAAGTAACCTCAGGCCCCCGCAGGCCCAAGCAGGCTCCTGTATGCTAGCCTT 1440
Qy 1483 CACCTTTCGAGGCTGTCTGTGAACCTGCAAGGGGGAGGTGAGCCCTGCAACGTAACCTTG 1542
Db 1441 CACCTTTCAGAGGTGTCTCACTCACTCAAGGAGAGGCGGCCCAAGCTGAGCCTTG 1500
Qy 1543 TCTCCCAACCTGAGCCCCGACACCAAGAGGCTCTCCCTGTATCAAGCCGCTGTGAAT 1602
Db 1501 TCTCCCAACCTGAGCTCTGACCAACAGAGGCTCTCCCTGTATGAGCCGCTCTGTAAC 1560
Qy 1603 GACCAAGGCTGCCCCCGGAGCGAGCGATGAGATGAGAGGCCGAGAGAAAGATTTCCAACT 1662
Db 1561 GACCAAGGCTGCCCCCGGAGCGAGCGAGAGAGAGAGAGGCGCGAGAAAGATTTCCAACT 1620
Qy 1663 GATTAAGCGGTACTTCTAGACTTAAGAGAGTGTCTTACCAAGGAGGAACTATCTGAAGAT 1722
Db 1621 GACAAAGGCTATCTAATCTGTGAAGAGTCTCCACACAGAGGCGACGTACTGAAGAGC 1680
Qy 1723 CTCGAAATTATCATCTTGTGTGTTTCAAGACAGTGAAGAAAGAGACCCATGCGGAA 1782
Db 1681 CTGAGAGCTATGCTCTGTGTGTTTCAAGACAGTCAAGCAAGAGAGACTCAATGCCGAA 1740
Qy 1783 GCACTGAAAAAGTCTATATTCGGAATTTTGAACCTTTGCACAAAATTTCTACTAATTTT 1842
Db 1741 GCCTTGAAAGTCTCATATTCCTCAAAATTTTGAACCTTTGCACAAAAGTTCCAAATTTT 1800

Qy 1843 CTCAAGAAAAATTGACCAACGACTTGCCCTGTGGGAAAGGCGGCTCAATATGCCCAATCA-- 1900
Db 1801 CTGAAGAAAAATTGACCAACGACTTGCCCTGTGGGAAAGGCGGCTCAATATGCCCAATCA 1860
Qy 1901 -GAGATTAACAAAGATGCGGATGTCTAGTGAAGAACTTTAGGGGATGAAGACCTG 1959
Db 1861 GGAAGATTAACAAAGATGCGGATGTCTAGTGAAGAACTTTAGGGGATGAAGACCTG 1920
Qy 1960 GCGGCTCACTGTGGAACACAGAGGCTTTGAGAGGCTCGAGAAATGGAATCAAGAGC 2019
Db 1921 GCGGCTCACTGTGGAACACAGAGGCTTTGAGAGGCTCGAGAAATGGAATCAAGAGC 1980
Qy 2020 TCCCGGCGGCTGGAAGACTTCTGACAGACTTTGAGTGTGAGAGGTTGTTACTTACG 2079
Db 1981 TCGCGGCGGCTGGAAGACTTCTGCGCGAGACTTTCAGAGCTGCAAGAGTGTGTTACTTCCG 2040
Qy 2080 CTCAACACTTCTCTCTGCGGCACTGCAACCGGCTCATGCACTACAGCAAGTCTTGAG 2139
Db 2041 CTCAACACTTCTCTCTGCGGCACTGCAACCGGCTCATGCACTATTAACATGTCTTGAG 2100
Qy 2140 CGGCTGTGCAACACACCCCGGAGCGAGCCGACTTCAGGAGATGCGGAGCGCTTGG 2199
Db 2101 AGGCTGTGCAACACACCCCGGAGCGAGCCGACTTCAGGAGATGCAAGACTGCGCTG 2160
Qy 2200 GCAGAGATCAAGGAGATGTGGCAAGCTCCACGCTACAGATGATCAAGATGAGAAATTC 2259
Db 2161 GCGAGATCAAGGAGATGTGGCAAGCTGCAAGTGTATCATATGATCAAGATGAGAAATTC 2220
Qy 2260 CAGAACTGTGCAAGACTCAAGAAAGATTTGATGGCAATTGACATCTTGTGTTCCGGGA 2319
Db 2221 CAGAACTGTGCAAGACTCAAGAAAGATCTGATCGGCAATTGACAACTTGTGATTCAGAGA 2280
Qy 2320 AGGAGTTCATCCGCTGAGGAGGCTCTGAGCAAGCTCTGAGGAGAGGGGCTCCAGAGGCG 2379
Db 2281 AGGAGTTCATCCGCTGAGGAGGCTCTGAGCAAGCTCTGAGGAGAGGGGCTTCAGAGGCG 2340
Qy 2380 ATGTTCTTCTGTTCAGAGAGTCTCTGTATATACAGACCGGGGCTGACGCGCTTCAAT 2439
Db 2341 ATGTTCTTCTGTTCAGAGAGTCTCTGTATATACAGACCGGGGCTGACGCGCATCTAAT 2400
Qy 2440 CAGTTTAAAGTTCACGGGAGCTCCGCTCTATGAGCATGAGATTTGAGAGAGCGAAGAC 2499
Db 2401 CAGTTTAAAGTTCACGGGAGCTCCGCTCTATGAGCATGAGATTTGAGAGAGGTAGAGAG 2460
Qy 2500 GAGTGGGGGGTGGCCCACTGACTGACGCTCGGGGGGCAAGCGGAGTCCATCATGCTGAGC 2559
Db 2461 GAGTGGGGGGTGGCCCACTGACTGACGCTCGGGGGGCAAGCGGAGTCTATCATGCTGAGC 2520
Qy 2560 GCGAGTTCCTGCTCGAGATGAGAAAGTGGTTGAGACATTCAGATGAGCAATTGACCTG 2619
Db 2521 GCGAGTTCCTGCTCGAGATGAGAAAGTGGTTGAGACATTCAGATGAGCAATTGACCTG 2580
Qy 2620 GCGGAGAAAGACAGAGGCCCCGCTGAGTTCTGTGCGCAGAGGCCCCCTGCAACAAG 2679
Db 2581 GCGAGAAAGACAGAGGCCCCGCTGAGTTCTGTGCGCAGAGGCCCCCTGCAACAAG 2640
Qy 2680 TCCCGTGAAGAGCAAGCGGGCTGACAGAGGTGAGAGATGACTGAGAGGCTCGGCGC 2739
Db 2641 TCCCGTGAAGAGCAAGCGGGCTGACAGAGGTGAGAGATGACTGAGAGGCTCGGCGC 2700
Qy 2740 ACATGCTGAGAGCGCCAGGCCCCGACCGGCAACAATGTTGCAAGTGTGCTGAGC 2799
Db 2701 ACATGCTGAGAGCGCCAGGCCCCGACCGTGGCAACAAGATGTTGCAAGTGTGCTGAGC 2760
Qy 2800 GCGAACACACGAGCTTCCATGTGAGACTTCAGACTGCAAGTGGAGAAATCAATGTTGCTGA 2859
Db 2761 GCGAACACACGAGTCTCCATGTGAGACTTCAGACTGCAAGTGGAGAAACAGTCTCTGAG 2820
Qy 2860 AACCTGTGAGGAATTTCAAAACAGCAACGGGTGGCAGAGCTGTGGGTGTTTCA 2919
Db 2821 AACCTGTGAGGAATTTCAAGAACAGCAATGAGCTGTGGAGAGCTGTGGGTGTTTCA 2880

Oy	1030	CCCGCTCTCTTTAGCCGGGGGTCATCATTTGGGTTCACTGGTCCGACTCAGAACAGGTT	1089
Db	1129	GCTGCTTCTTTCAGCCGAGGCTCCCTCTTCAGATACAGTGGAAACCTCAGAAACAATA	1188
Oy	1090	CTCGACTATGTTAAAGAGAGACATATAAGAGGCGCTTTGAAGAGAACACAGCAAG	1149
Db	1189	GTAGTTATTTGTCMAAGACGCTGGATGAAAGAAATTCATACGAAAGCGGCACTGTAAG	1248
Oy	1150	ATTCC--ATTCTATCCGAGGCTTGCTTACAGCCTACAGAACTGMAATTCGGAAAGTCGTG	1206
Db	1249	ACTGGGACATCTCTTCATGCTCTGACTGTAGTCTGCTTAACAGAGGCTCTCTTCAACC	1308
Oy	1207	GAGCAGTCTCAGCAGAGACCAACGCTTTCATTTGGAGAAGTCCGAATCTCCAGGGGGC	1266
Db	1309	GATGGCTTGAGAACTTCTGCTCCCTCTGCTTCAGCAATATGCTCTTTTATATCC-----	1361
Oy	1267	CAGAGCTGCGGGCAGAAAGAAACCGAAGTTTCCGCGGGGAGCCGGGGTGGCAACCG	1326
Db	1362	--ACCCCTTAGTTCTTTTGTCCCTCCCTGCTGCTGCCAATTTGAAGACACACAGAG	1418
Oy	1327	AGCCCTGCGCGAGAGAGAGCCCGCGGGTAAACAAGCGCGAGCGGAGCCGCTCGAGC	1386
Db	1419	CTCCCTTGATGATCCGAGGCTCCCGTATCAAGAGCACACAGACAGAGAGACAGAGG	1478
Oy	1387	CCCAACGAGAAAGAGAGAGGTGTTAAGATAGAACCCAGACAGATTAACCTCAGCCC	1446
Db	1479	ACCA--TCATCATCTGACGGCCCCAGCACAGTCGGCCCATCTCCCTGAGCCCCCGTG	1536
Oy	1447	CCGACGCCAACACAGAGGTCCTCTGACTGGCACTGCTTACCTTCCGAGCTCTGTAGAC	1506
Db	1537	CTCCGGCTGATCCAGGCTTTTCTATGATATGCTCTCAGCCTTCTCCCTCAGCCTAAG	1596
Oy	1507	TCCGAGGGGGAGTGGCCCCCTTGCCACAAGTACTTGTCTCCAACTGAGCCCCGAGAAC	1566
Db	1597	AGCCACCTGAGACCTGTGTGCTTGAAGTTTCAAGCCCG-----CTGATGACGCTGAG	1647
Oy	1567	AAGCAGGCTCTCCCTTGATCAGCCCGCTGTGAATGACAGGCTCCGCCCGGAGCGAC	1626
Db	1648	CAGGATGATCCCTGCTCAGCCCTGTGCTCAGTGGCGCTGGTACAGCCAGATGAC	1707
Oy	1627	GATAGAGATGAGGCGCGAGAGAGATTTCCCACTGATTAAGGTACTTCAATGCTAAG	1686
Db	1708	AACCAAGAGGAGCAAAACACAAACACATGCGCAAGATGAGGCTATTTCAATAGCCAAG	1767
Oy	1687	GAACTGCTACACACGAGCGAAACAATCTGAAAGATCTCGAAGTTATCACTTCGTGTTT	1746
Db	1768	GAGATCTCGCTACAGAACGAACCTATCTGAAAGATTTAGAAATTAACACTGTGTGTTT	1827
Oy	1747	CAGAGCAGTGAACAAAGAGACGCCATGCGGAGACACTGAAAGTCTCATATTTCCG	1806
Db	1828	CGAGCGGTGATCAAGAGAGAGGCCATGCTCAGACCTGATGGCCGTGCTTTCTCC	1887
Oy	1807	AATTTTGAACCTTTGACAAATTTCAATCTAATTTTCTCAAGAAATTGAGCAACGACTT	1866
Db	1888	AACATTATTCAGTCTACAGATTTCCACAGAGCTTCTTCAAGGTGAGAACAGGCTG	1947
Oy	1867	GGCCTGAGGAGGCGGCTCAAAATGCCA---AATCAGAGATTACAAAGATGGGGAT	1923
Db	1948	GCACTCTGGAGAGGCGCTTCCAGTGCCCACTTAAGATGATACACAGCGATGGGGAC	2007
Oy	1924	GTCATGTGTAAGAACTTTCAGGGCATGAGCACTGGCGGCTCACTGTGGAAACAGC	1983
Db	2008	ATCTCTCTCAGAAACATGCGTCAGTTAAGGAATTTACTAGCTACTTCAAAAGACGAT	2067
Oy	1984	GAGGCTTGAAGCGCTTGAAGATAGGAATCAAGAGCTCCGGGCGCTGAGAACTTCTGC	2043
Db	2068	GAGGTCTTAACAGAACTGGAAGAGCCCAAAACACTGTAAAAAGCTGAGGACGTATAC	2127
Oy	2044	AGAGACTTTGAGCTGACAGAGGTGTGTTACTTACCGCTCAACACTTCTCTGCGGCA	2103
Db	2128	AAAGATTTGAGCTTCAAAAGGTCTGCTACTGCTCTCAACACATTTCTGCTGAGGCC	2187
Oy	2104	CTGACCGGCTCATGCACTACAGAGGTCTGTGAGCGGCTGTGCAACACACCCCGCG	2163

Db	2188	GTCCAGAGCTAAGTCACTACCGTGTGCTGCTAGACCCGCTGTGTGCTCACTTCTCT	2247
Qy	2164	AGCCACGCCGCACTTCAGGGGACTGCGGAGCCGCTTGGCAGAGATCACGGAATGGTGCA	2222
Db	2248	GGGCAACCCGCACTATGCGGACTGCGCATGAGGCACTGAAGGCCATACAGAAAGTACACCC	2307
Qy	2224	CAGCTCCACGGTACGATGATCAAGTGAAGATTTTCCAGAACTGCAAGAACTCAAGAA	2283
Db	2308	GAGCTCCAGCAAAAGCTTATCCCGGCTGGAAAACTACAGAAATTGACGAGCTACAGCGA	2367
Qy	2264	GATTTGATTTGGATGAAATCTTGTGTGTTCCGGGAAGGGAGTTCACTCGTGGGACG	2344
Db	2368	GACCTGTGCGGTGTAAGAAACCTCATTTCTCTCGGAGAGGATTTATCGTGAAGGCTGC	2427
Qy	2344	CTCAGCAAGCTCTCGGGGAGGGGCTCCAGACGCGCATGTTCTTCTGTTCAACACGTC	2403
Db	2428	CTGCAAAAGCTCACCAAGAGGGCTCGAACAAGAGATGTTTTTTCTGTCTCAGATATG	2487
Qy	2404	CTGCTATACAGAGCCGGGGGCTGACGGCTCCATTCAGTTTAAAGTCCAGGGCAGCTC	2463
Db	2488	TTACTGTATACAGAAAGATGTCAAGAGGCAAGTCAATTCGGATCCGTGGCTTCTT	2547
Qy	2464	CCGCTCTATGGCATACGATTTGAAGAGCCAAAGACGATGGGGGTGCCCATCTGCTG	2522
Db	2548	CCACTCCGTGCATCTGTGTGAAGAAAGTGAATGATGATGTTCTTTCATTGCTTC	2607
Qy	2524	ACCCCTCGGGGGCAGCGGCAAGTCCATCATCGTGGCCGCAAGTTCGGTCCGAGTGAAG	2583
Db	2608	ACCATCTATGCAAGCTCAGAAAACAATTTGTGATGACGCACTCGCTTGAAAGAGAA	2667
Qy	2584	AAGTGGTGTGAAGACATCCAGATGGCCATTGACCTTGGCGGAAGAGCAGCAGCCCGCC	2643
Db	2668	AAGTGAATGCAAGAACCTGAATGCAAGCATCCAGAGGCAAGACATATGATGATCACTC	2727
Qy	2644	CCTGATTTCTGGCCAGCAGCCCCCTGACAACAAGTCCCTGATGAAGCCACGGGCT	2703
Db	2728	CCAGTGTGCTGG---GAGGCCCGGTGTATCTGTACCTCTAGATCTTCTGATGAATCTC	2784
Qy	2704	GACCAAGAGTCAAGGAATGACTGAGGCGCTCGGCAATCGCTGGAAGCGCAGGCCCCG	2765
Db	2785	TCTTGGAAATTCAGAAATGATGTCAGAGAAACCGGGGCTCCCTGGAAGGGAACAGCTG	2844
Qy	2764	CACCGCGCAACACAAATGATGTCAGTGTGTGGAACCTGCTGAGGAAATTCAAAAC	2823
Db	2845	CACCGGGCAATACAAATGCAATGTGTGTGCTGTATACGTAATACAAATGTGTCCAGAC	2904
Qy	2824	GACTTCAGCATCGAGTGAAGATCAAGTTGTGTGAAACCTGCTGAGGAAATTCAAAAC	2883
Db	2905	GACCAAGTGAAGCTGTGGAACACAGCTTTCAGAGATCTGTGAGAGAAATTCAGAAAC	2964
Qy	2884	AGCAACGGGTGGCAAGACGTGGTGTGTTCAAACTTCGCGCTTCTTCTTCAACAA	2943
Db	2965	AGTAATGGCTGCAAGAACTGTGGGTGTCTTTCACAACTTCTGTGTCTTCTTATATA	3024
Qy	2944	TCACACCAAGCAATCATCCCTTGGCCAGCTGCTCTGCTCGGCTACTCGCTCACATC	3003
Db	3025	ACAATATGAGATGATACCCCTTGGCCAGCTTCCCATATTTGGGTTCAGGGTGAAGCTC	3084
Qy	3004	CCCTTGAAGTCCGAAGCATCCAGAAAGATTAAGTTCAGGTGCACTTCAAGTCCAC	3063
Db	3085	CCAGAGGAGGCTGACAGCATCCAAAGACTATGTCTTCAAGCTCCAAATTCATCAC	3144
Qy	3064	GTCATCTACTTCAGAGGCGGAAGAGAGTACAGTTCGAAAGAGTGAAGTATATCCG	3123
Db	3145	GTCATCTCTTCGCGGCTGAAGAGCAATTCATTTGAAGGTGAAGTCAATCAAA	3204
Qy	3124	AGTGCACACAGCTGTGCTCGCGACCCCACTGTGTGAAGCCAAAGATCTCTTGTGTAT	3183
Db	3205	AGGGCGAGCACTCACAGGGAAGACCCCAAGTTTCACTAGAGCTGTCAATCATCTCT	3264
Qy	3184	TGATGGCGGACACATCTGTTTCCGAGTGGCTGCTTCC	3223

Db 583 ATGCGAGATTATGTAAGAAACCTTGATCGAAGAACCTCAAGCAATGATACTTGCC 642
Qy 610 CACGAGAACCACTAGAGAGCAAAATTCGTGGAATTTCAACATAACCAATTTGACAAACA 669
Db 643 AACGAGAGAAATCTTAGAAAGAAATCTAGACTTCATCAGAGAGCACAGGCGCAAGACT 702
Qy 670 CCAGCAAAATCAGATTTTCAAGCTCCTAGAGATTTGCCGTGAGATGATGAAATC 729
Db 703 CCGGCAAGTCAGATTTTCAAGGTCCTTAGATTTGCAAGGAAGCTGGAATGTATGCAATC 762
Qy 730 CGGTTCAGCCGCGCAAGAGCAAGGAAGCAAGAAATCAATCTGCGCTTGCACACG 789
Db 763 AGGTTTCACTGGCTTCTGACAGAAAGGAGCAAGATTAATCTAGCAGTTTCTCAGATG 822
Qy 790 GGAATTTAGTGTTCAGGGTTTCACTAAGATCAATGCTTCAACTGCGGCAAGTCCG 849
Db 823 GGTGTCTGTGTTCAGGGTACCAACAAATCAATCTTCAACTGCTGCAAGTCCGT 882
Qy 850 AAGCTGAGCTTCAAGAGAGAGCGCTTCTCATCAAGCTCCGCGCAATGCAATAGTGC 909
Db 883 AAATTAAGCTTCAAGAGAAAGATTTCTTATCAAACTCCACCTGAGTCCATGGGCC 942
Qy 910 TACCAAGATACCTTGAATTCCTGATGCGCAGTCGGGATTTCTGCAATCTTCTGAAA 969
Db 943 TACCAAGACACGTTAGAGTTCTTGTTGGTAGCAGAGATGAATGTAAAGACTTGGAG 1002
Qy 970 ATCTGTGTGAACATCATGCTTCTTTAGACTTTTGAAGCCCAACCAAGCCCAAG 1029
Db 1003 ATATGTGTGAGTACCAATACCTTTTGAAGCTCTGACCAAGCTTAAAGCCAAAGCAAG 1062
Qy 1030 CCGGCTCTCTTAAAGCGGGGATCATATTCGGTTCAGTGTGAGACTCGAAGCAGGTT 1089
Db 1063 GCTGTCTTCTTCAAGCGAGGCTCTCTTCAAGATCAGTGAAGAACTCAAGAAACAATCA 1122
Qy 1090 CTCGACTATGTAAAGAGAGAGCAATAGAGAGTGCATTTGAAGAAAGCAAGCAAG 1149
Db 1123 GTAGATTAATGTAAAGACGCTGGAATGAAGAAATTCATACGAAGACGGCAAGTAAG 1182
Qy 1150 ATTCT--ATTCTATCCGAGGCTTGTCTTCAAGCTTACAGATCAAGATTTGGAAATGCTG 1206
Db 1183 ACTCGAGCATCTTTCATGCTCTGACGTGATCTGCTAAACAGACGCTCTCTTCAACC 1242
Qy 1207 GAGCAGTCTCAGAGAGACCAAGCTTATCAATTTGAGAGAGTGCAGATCTCCAGGGGAG 1266
Db 1243 GATGCTTGAAGACTTCTGCTCTCTCTGTCTTCAAGCAAAATGTCTCTTTATTC----- 1295
Qy 1267 CAGAGCTGCGGCGAGAGAAAGAACCGAAGTTCGCGCGGAGAGCGGGGTGCAACCG 1326
Db 1296 ---ACCCCTTAATTTCTTGTCTTGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1352
Qy 1327 AGCCCTGCGCGAGAGAGAGGCTTGAAGATAGAGCCAGAGAGTAAACCTCAGGCC 1386
Db 1353 CTCCTTGTGATCCCGAGGCTCCCGTCAACAGAGACAGCAGAGAGAGAGAGCGG 1412
Qy 1387 CCCAGGAGAGAGAGAGAGGATCGTTAAGATAGAGCCAGAGAGTAAACCTCAGGCC 1446
Db 1413 ACCA--TCACTATCTAGCGGCCCAAGACAGCTCGGCCATCTCTCTGAGCCCGCGT 1470
Qy 1447 CCGCAGCCAAAGCAGAGCTCTTGAATGAGAGTCTTCACTTTTCAAGCTGTCTGTGAAC 1506
Db 1471 CTCGCGCTGTGTCAGAGCTTTTCTATGGAATGCTCTCAGCTTCTCCCTCAGCCTGAAG 1530
Qy 1507 TCGCAGGGGAGAGTGGCTTCTGCAAGAGTGAATCTTCTCCAACTGAGCCCGAGACC 1566
Db 1531 AGCCACTGAGCTGTGCTCTGAGCTTCAAGCGCG-----CTGAATACAGCTGAG 1581
Qy 1567 AAGCAGGCTCTCTCTTGAATCAGCCGCTGTGAATGACAGGCTGCGCCCGCAGAG 1626
Db 1582 CAGGCTGCACTCCCTGTGTCTCAGCCCTGTGTCTCAGTGGCGCTGTACAGCAGAGATGAGC 1641
Qy 1627 GATGAGAGTGAAGGCGCGAGAGAAAGATTTCCAACTGATTAAGCTTCAATAGCTAAG 1686
Db 1642 AACCAAGAGAGAGAGAAACAAAGACATGCTCAGAAAGATGAGGCTTATTTCTATGCAAG 1701

Qy 1687 GAAGTGTCTAACCAAGAGAGACATATCTGAAGGATCTGAAAGTTATCACTTGTGTTT 1746
Db 1702 GAGATTCGTGCTTACAGAGAGACCTATCTGAAGGATTTGAAAGTATCACTGTGTGTTT 1761
Qy 1747 CAGAGCAGATGAGCAAGAGAGAGAGCGCATGCGGAGAGCACTGAAGTCTCATATTTCCG 1806
Db 1762 CGAGGCTGTGTATCAAGGAGAGAGGCGCATGCTGAGCCCTGATGGCCGCTTTCTTC 1821
Qy 1807 AATTTGAACCTTTGCACAAAATTTCAATCTAATTTTCTAAGGAAATTTGAGACAGACTT 1866
Db 1822 AACATTAATCTAGTCTACAGATTCAGAGGCTTCTTCAAGAGGTGAAACAGAGCTG 1881
Qy 1867 GCGCTGAGAGAGGCGCTCAATGCGCA--AATCAGAGATTAACCAAGAAATCGCGAT 1923
Db 1882 GACTCTGAGAGAGGCGCTTCAAGTGTCTTCACTTAAAGTATACAGAGATCGGGAGC 1941
Qy 1924 GTCATCTGAGAGAACTTCAAGGAGATGAAGCACTTGCGGCTTCACTGTGAGAGCAGC 1983
Db 1942 ATCTCTCTCAGAGAACTGCGTCACTTAAGGAATTTACTAGCTATCCAAAGACAGAT 2001
Qy 1984 GAGGCTTGAAGGCTTGAAGATGAATCAAGAGCTCCCGCGCTGAGAACTTCTGC 2043
Db 2002 GAGTCTTAACAGAACTGGAAGAGCCACAAACACTGTAAAGCTGAGGCGAGCTTAC 2061
Qy 2044 AAGAGCTTGAAGTGAAGAGGTGTCTTACAGCGCTCAACCTTCCCTGCGGCGCA 2103
Db 2062 AAGAGTTGAAGTCCAAAGGCTGTCTACCTGCTTCAACATCTTCTGTAAGCCC 2121
Qy 2104 CTGCAAGGCTCAGTCACTTCAAGAGGCTCTGAGAGGCTGTGCAACCAAGCCGCG 2163
Db 2122 GTCAAGAGGCTAGTCACTACGCTGTGCTGAGGCGGCTGTGCTCACTACTCTCT 2181
Qy 2164 AGCCAGCCGACTTCAAGGAGCTGCGAGCGCTTTTGGCAGAGATCAAGAGATGTGCA 2223
Db 2182 GGGCAGCGGAGACTATGCTCCAGCTGAGGAGCATGMAAGGCATCAAGAGATGACAC 2241
Qy 2224 CAGCTCAGGATCAATGATCAAGAGGATTTTCAAGCTGACAGATCAAGAA 2283
Db 2242 GAGCTCAGCAAGAGCTTACCGGCTGAGAAACCTACAGAAATTTGAGAGCTCAGAG 2301
Qy 2284 GATTTGATTTGCAATGATCAATCTTGTGTTCGAGAGAGGATTCATCCGTCTGAGCAG 2343
Db 2302 GACCTGTGTGTATGAAGAAACCTCATTTGCTCTGAGAGGAGATTAATCCGTGAGGCTG 2361
Qy 2344 CTCAGAGCTCTCGGAGAGAGGAGCTCAAGAGCGCATTTCTTCTTTCAAGCAGTTC 2403
Db 2362 CTGCAAGAGCTCAACMAAGAGGCGCTGCAAGAGGATTTTTTCTGTCTCAGATATG 2421
Qy 2404 CTGCTATTAAGAGCGGGGCTGACGCGCTCAATCAATTTAAAGTCAAGGCGCAGCTC 2463
Db 2422 TTGCTGTATTAAGAGAAAGTGTCAAGAGGCGAGTCAATTTCCGATCCGTGCTTCTT 2481
Qy 2464 CCGCTTATGCAATGCAATTTGAGAGAGAGAGAGAGAGTGGGGGTGCGCCACTGCTG 2523
Db 2482 CCACTCCGTGAGATGTGTGATGAAGAAAGTGAAGATGTGTCTCTCATTTGCTTC 2541
Qy 2524 ACCCTCGGAGCGAGCGCAATCATCTGTGCGCGCAATTTCTGCTTCCAGATGAG 2583
Db 2542 ACCATTAATGCAAGCTCAAGAAACAATTTGTGTAGCAGCAGAGCTCGGCTAAGAAAGGAA 2601
Qy 2584 AAGTGTGTTGAAGATCAATCAAGATGAGCTTGAACGTGCGGAGAGAGAGAGCGCGCG 2643
Db 2602 AAGTGTATCAGAGACTTAATGAGAAATTCAGAGAGCTATTCGATCTTACCTC 2661
Qy 2644 CCTGAGTTCTGAGCAGAGCGCCCTCTGACAAAGATCCCTGATGAAGCCACCGGAGCT 2703
Db 2662 CAGGTCTGTGAGAGGCGCGGTATTACTGTGACCCCTAGATCTTGTATGAAGTCTCT 2721
Qy 2704 GACAGAGATCAAGAGATCACTGAGCGCTTGCAGCATGCTGAGAGCGCAGGCTCCG 2763
Db 2722 CTGGAAGAAATCAAGAGATGTGAGAGAAAGCGGGGCTGCTGTGAGGGGAAAGCCAG 2781

```

QY 2764 CACCGCG--CAACAATGTCAGTGTGTGTCAGCCGCAACACAGCCTGTCCATG 2821
DB 2782 CACCGGGGCCCAATACAAATGATGTGTGTGTGTAACGTAATACAAATGTGTCCAGG 2841
QY 2822 TGGACTTCAGCATTCGAGTGTGAGAAATCAGTGTGTGAAAACCTGTGAGAAATTCAAA 2881
DB 2842 CAGACCAAGTCAGCTGTGTGAGAACACAGCTTTCAGAGATATCTGTGAGAAAGTTCAGA 2901
QY 2882 ACAGCAAGGGGTGGGAGAAAGCTGTGTGTGTGTGCAAACTTGTGCGCTGTCTTCTTACA 2941
DB 2902 ACAGTAATAGCTGGCAAGAAAGCTGTGTGTGTGTGCAAACTTGTGTGTGTGTGTATTA 2961
QY 2942 AATCAACACCAAGCAATATCTCCCTTCCAGCCTGTGTGTGTGTGTGTGTGTGTGTGTGT 3001
DB 2962 AAACATCATAGATGATCTACCCCTGTGGCCAGCTCCCACTATTGGTTTACAGGTGAGCC 3021
QY 3002 TCCCTCTGATGTCGAGAAACATTCAGAAAGATCTAGTGTTCAGCTGCACTTCAAGTCCC 3061
DB 3022 TCCCCAGGAGGGGTGACAGATCCCAAAAGACTATGTCTTCAAGCTCCAAATTCATCTC 3081
QY 3062 AGCTACTACTTTCAGAGGGGGAAGGAGTACAGTTCGAAAGGTGATGAGTATCC 3121
DB 3082 AGCTACTACTTTCAGAGGGGGAAGGAGTACAGTTCGAAAGGTGATGAGTATCC 3141
QY 3122 GAGTGCACCAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3181
DB 3142 AAAGGCGCAGAGCTACACAGGAGAGCCCAAGTTTCACTAGAGCTGTGTGTGTGTGTGT 3201
QY 3182 ATTGATGCGCGGACACACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3223
DB 3202 CTCAGAGGGGTGAGGAGAGATCAGAAAGAGGAGCATGCC 3243

```

RESULT 7
BX364606

LOCUS BX364606 968 bp mRNA linear EST 08-APR-2004

DEFINITION BX364606 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens

ACCESSION CDNA clone CS0DC027YA10 5-PRIME, mRNA sequence.

VERSION BX364606

KEYWORDS BX364606.2 GI:46289717

SOURCE EST.

ORGANISM Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo.

1 (bases 1 to 968)

L1,W.B., Gruber,C., Jessee,J. and Polyes,D.

Full-length cDNA libraries and normalization

Unpublished (2001)

On May 5, 2003 this sequence version replaced gi:30370819.

Contact: Genoscope

Genoscope - Centre National de Sequencage

2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime

end enriched, double-strand cDNA was digested with Not I and cloned

into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library

was normalized. Library was constructed by Life Technologies, a

division of Invitrogen. This sequence belongs to sequence cluster

7229.f

For more information about this cluster, see

http://www.genoscope.cns.fr/cdna?c=CS1AC007ZE06P1c=7229.f.

FEATURES

source

1.968

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CS0DC027YA10"

/rname="NEUROBLASTOMA COT 25-NORMALIZED"

/clone_id="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"

/note="1st strand cDNA was primed with a NotI-oligo(dT)

primer. Five prime end enriched, double-strand cDNA was

ORIGIN

digested with Not I and cloned into the Not I and EcoR V

sites of the pCMVSPORT 6 vector. Library was normalized."

```

QY 1110 AGGATATAGAGAGTGCAGTTTGAAGAGAGACAGAAAGATTCATTCGAGAGCCT 1169
DB 1 AGGATATAGAGAGTGCAGTTTGAAGAGAGACAGAAAGATTCATTCGAGAGCCT 60
QY 1170 TGTCTCAGAGCTTACAGAACTGAATTCGAAATGCTGTGAGAGCTCAGCAGAGCAG 1229
DB 61 TGTCTCAGAGCTTACAGAACTGAATTCGAAATGCTGTGAGAGCTCAGCAGAGCAG 120
QY 1230 CTTTACATTTGAGAGAGTGTCCGAATCTCCAGAGGGGCGCAGAGCTCCGCGGAGAAAG 1289
DB 121 CTTTACATTTGAGAGAGTGTCCGAATCTCCAGAGGGGCGCAGAGCTCCGCGGAGAAAG 180
QY 1290 ACCGAAAGTTTCCGCGGGGAGCCGGGGTCCGACCCGAGCCCTGCGCGGAGAGAGCC 1349
DB 181 ACCGAA-CTTTCGCGGGGAGCCGGGGTCCGACCCGAGCCCTGCGCGGAGAGAGCC 239
QY 1350 CGCGGGTAAACAAGCAGAGCGGAGCAGAGCCGCTCGGCGCCACAGAGAGAGAGAGGT 1409
DB 240 CGCGGGTAAACAAGCAGAGCGGAGCAGAGCCGCTCGGCGCCACAGAGAGAGAGAGGT 299
QY 1410 CGTTAAGATAGAGACCCAGAGAGTAACTTCAGGCCCCCGCAGAGCAGAGGCTCCCT 1469
DB 300 CGTTAAGATAGAGACCCAGAGAGTAACTTCAGGCCCCCGCAGAGCAGAGGCTCCCT 359
QY 1470 GACTGGAGTCCCTCACTTTCGAGAGCTGTGTGTAACTCGAGGGGGAGTGGCCCTGC 1529
DB 360 GACTGGAGTCCCTCACTTTCGAGAGCTGTGTGTAACTCGAGGGGGAGTGGCCCTGC 419
QY 1530 CAACGTGACCTTGTCTCCCAACTGAGGCCCGCAGACCAAGAGCGCTCTCTTGATCAG 1589
DB 420 CAACGTGACCTTGTCTCCCAACTGAGGCCCGCAGACCAAGAGCGCTCTCTTGATCAG 479
QY 1590 CCGCGTGTGAATGACAGAGGCTGCCCCCGGAGCAGATGAGATGAGGCGCGAGAGAA 1649
DB 480 CCGCGTGTGAATGACAGAGGCTGCCCCCGGAGCAGATGAGATGAGGCGCGAGAGAA 539
QY 1650 GAGATTTCCCACTGATTAAGGCTACTTATAGCTAAGGAAGTGTTCACACGAGCGAAC 1709
DB 540 GAGATTTCCCACTGATTAAGGCTACTTATAGCTAAGGAAGTGTTCACACGAGCGAAC 599
QY 1710 ATATCTGAAGATCTGAAGTATCATCTTGTGTGTTCAGAGACAGTGAAGAAAGAGA 1769
DB 600 ATATCTGAAGATCTGAAGTATCATCTTGTGTGTTCAGAGACAGTGAAGAAAGAGA 659
QY 1770 CGCCATGCGCGAAGACAGTGAAGTCTCATATTCGCCGAATTTGAACCTTTGCAAAATT 1829
DB 660 CGCCATGCGCGAAGACAGTGAAGTCTCATATTCGCCGAATTTGAACCTTTGCAAAATT 719
QY 1830 TCATCTAATTTTTCAGAGAAATGAGCAAGACTTCCCTGTGTGAGAGGCGCTCAAA 1889
DB 720 TCATCTAATTTTTCAGAGAAATGAGCAAGACTTCCCTGTGTGAGAGGCGCTCAAA 779
QY 1890 TGCCCAATCAGAGATTCAGAAAGATGCGGATGCTCATGTAAGAACATTGAGGGCAT 1949
DB 780 TGCCCAATCAGAGATTCAGAAAGATGCGGATGCTCATGTAAGAACATTGAGGGCAT 839
QY 1950 GAAAGACCTGCGGCTGCTCTGTGAGACACAGAGGCTTTGAGGCGCTGGAAGATG 2009
DB 840 GAAAGACCTGCGGCTGCTCTGTGAGACACAGAGGCTTTGAGGCGCTGGAAGATG 899
QY 2010 AATCAAGAGCTCCGCGGCTGAGAACTTCTGAGAGACTTTGAGCTGCAAGAGTGTG 2069
DB 900 AATCAAGAGCTCCGCGGCTGAGAACTTCTGAGAGACTTTGAGCTGCAAGAGTGTG 959
QY 2070 TTACTTACC 2078

```

```

Db          960 TTACTTACC 968

RESULT 8
BX354921
LOCUS      BX354921 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
DEFINITION CDNA clone CS0DC027YA10 5-PRIME, mRNA sequence.
ACCESSION  BX354921
VERSION     BX354921.2 GI:46553760
KEYWORDS   EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Homiidae; Homo.
REFERENCE  1 (bases 1 to 1000)
AUTHORS   Li, W.-B., Gruber, C., Jessee, J., and Polayes, D.
TITLE     Full-length cDNA libraries and normalization
JOURNAL   Unpublished (2001)
COMMENT   On May 5, 2003 this sequence version replaced gi:30383921.
Contact: Genoscope - Centre National de Sequencage
          2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
          Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
          1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
          end enriched, double-strand cDNA was digested with Not I and cloned
          into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
          was normalized. Library was constructed by Life Technologies, a
          division of Invitrogen. This sequence belongs to sequence cluster
          7229.f
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?c=CS0DC027BA050P1&c=7229.f.
FEATURES
source
1.1000
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DC027YA10"
/issue_type="NEUROBLASTOMA COT 25-NORMALIZED"
/clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA
was digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN
Query Match 27.3%; Score 940.2; DB 4; Length 1000;
Best Local Similarity 99.2%; Pred. No.1.5e-240;
Matches 961; Conservative 4; Mismatches 2; Indels 2; Gaps 2;

QY 1110 AGGACATAAGAGAGTGCAGTTTGAAGAGACAGACAGATTCATTCTACCGAGCCT 1169
    |||||
    1 AGGACATAAGAGAGTGCAGTTTGAAGAGACAGACAGATTCATTCTACCGAGCCT 60
    |||||

QY 1170 TGGTTACAGCCTTACAGAACTGAATTCGGAAGTCTGAGACATCTTCAGAGACACCG 1229
    |||||
    61 TGGTTACAGCCTTACAGAACTGAATTCGGAAGTCTGAGACATCTTCAGAGACACCG 120
    |||||

QY 1230 CCTTACATTGAGAGAGTGCAGATCTCCAGAGGGGCGAGAGTGCCTGGGAGAGAAAGGA 1289
    |||||
    121 CCTTACATTGAGAGAGTGCAGATCTCCAGAGGGGCGAGAGTGCCTGGGAGAGAAAGGA 180
    |||||

QY 1290 ACCGAAAGTTTCGCGCGGGAGGCGGGGCTGCACCCGAGCCCTGCGCCGAGAGAAAGCC 1349
    |||||
    181 ACCGAAAGTTTCGCGCGGGAGGCGGGGCTGCACCCGAGCCCTGCGCCGAGAGAAAGCC 239
    |||||

QY 1350 CCGGGGTAAACAAGCAGCGGAGCGGCGCTCGGGCCCAAGAGAGAGAGAGAGGT 1409
    |||||
    240 CCGGGGTAAACAAGCAGCGGAGCGGCGCTCGGGCCCAAGAGAGAGAGAGAGGT 299
    |||||

QY 1410 CGTTAAGATAGAGACCAAGAGTAACCTCAGCCCCCGCAGCAAGCAGAGCTCCCT 1469
    |||||

```

```

Db          300 CGTTAAGATAGAGACCAAGTAACCTCAGCCCCCGGAGCAAGAGAGGCTCCCT 359
    |||||
QY 1470 GACTGCACTCTTCACTTCCGAGCTGTCTGTAACCTGCAAGGGGAGGTGCCCCCTGC 1529
    |||||
Db 360 GACTGCACTCTTCACTTCCGAGCTGTCTGTAACCTGCAAGGGGAGGTGCCCCCTGC 419
    |||||
QY 1530 CAACGAGACTTGTCTCCCAAGCTGAGCCCGGACACCAAGCAGGCTCTCCCTGATCAG 1589
    |||||
Db 420 CAACGAGACTTGTCTCCCAAGCTGAGCCCGGACACCAAGCAGGCTCTCCCTGATCAG 479
    |||||
QY 1590 CCGGCTGCTGAATGACACGAGCCTGCCCGGAGCGAGATGAGATGAGGGCCGGAGGA 1649
    |||||
Db 480 CCGGCTGCTGAATGACACGAGCCTGCCCGGAGCGAGATGAGATGAGGGCCGGAGGA 539
    |||||
QY 1650 GAGATCCCAACTGATTAAGCGTACTTCATAGCTTAAGAAAGTGTCTTACACCGAGGAAC 1709
    |||||
Db 540 GAGATCCCAACTGATTAAGCGTACTTCATAGCTTAAGAAAGTGTCTTACACCGAGGAAC 599
    |||||
QY 1710 ATATCTGAAGATCTTGAAGTATCACTTCGTGCTTTCAGACAGATGAGCAAGAGGA 1769
    |||||
Db 600 ATATCTGAAGATCTTGAAGTATCACTTCGTGCTTTCAGACAGATGAGCAAGAGGA 659
    |||||
QY 1770 CCGCATGCCGGAAGCACTGAAAAGTCTCATATTCGCCAATTTTGAACCTTTGACAAAT 1829
    |||||
Db 660 CCGCATGCCGGAAGCACTGAAAAGTCTCATATTCGCCAATTTTGAACCTTTGACAAAT 719
    |||||
QY 1830 TCATACTAATTTTCTCAGAGAAATTGAGACAGACTTCCCTGTGGAGGCCGCTCAAA 1889
    |||||
Db 720 TCATACTAATTTTCTCAGAGAAATTGAGACAGACTTCCCTGTGGAGGCCGCTCAAA 779
    |||||
QY 1890 TGCCCAATCAGAGATTCACAAAGAAATGGCGCATGCTGGAAGAAATTCAGGCAAT 1949
    |||||
Db 780 TGCCCAATCAGAGATTCACAAAGAAATGGCGCATGCTGGAAGAAATTCAGGCAAT 839
    |||||
QY 1950 GAAGCACTGGGCGGCTCACTGTGAGAGACAGACGAGGCC-TTGAAGGCCCTGAGAAATG 2008
    |||||
Db 840 GAAGCACTGGGCGGCTCACTGTGAGARCAAGCAGAGGCCCTTTGAGAGGCCGTGAGAAATG 899
    |||||
QY 2009 GAATCAAGAGCTCCGCGCGGCTGAGAGAACTTCTGCAAGAGACTTTGAGCTGAGAGAGTGT 2068
    |||||
Db 900 GAATCAAGAGCTCCGCGCGGCTGAGAGAACTTCTGCAAGAGACTTTTGTGAGAGAAAGTGT 959
    |||||
QY 2069 GTTACTTAC 2077
    |||||
Db 960 GTTACTTAC 968
    |||||

RESULT 9
BX354920/c
LOCUS      BX354920 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
DEFINITION CDNA clone CS0DC027YA10 3-PRIME, mRNA sequence.
ACCESSION  BX354920
VERSION     BX354920.2 GI:46552853
KEYWORDS   EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Homiidae; Homo.
REFERENCE  1 (bases 1 to 1072)
AUTHORS   Li, W.-B., Gruber, C., Jessee, J., and Polayes, D.
TITLE     Full-length cDNA libraries and normalization
JOURNAL   Unpublished (2001)
COMMENT   On May 5, 2003 this sequence version replaced gi:30381927.
Contact: Genoscope - Centre National de Sequencage
          2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
          Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
          1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
          end enriched, double-strand cDNA was digested with Not I and EcoR V
          sites of the pCMVSPORT 6 vector. Library

```


Db	892	TTCTTCCGTTCACAGAGCTGCTGCTATACACAGACCGGGGGCTGACGGCCCTCAATCAG	833
Qy	2443	TTTAAAGTCCACGGGCAAGCTCCCGCTCTATGAGCATGACGATTTAGAGAGCGAAGACGAG	2502
Db	832	TTTAAAGTCCACGGGCA-STCCCGCTCTATGGCATGACGATTTGAGGAAGCGAAGACGAG	774
Qy	2503	TGGGGGGGTGGCCCACTGCTGACCCCTCCGGGGCCAGCGGCAAGTCCATCATGCTGGCCGCC	2563
Db	773	T-TGGGGGTGGCCCACTGCTGACCCCTCCGGGGCCAGCGGCAAGTCCATCATGCTGGCCGCC	715
Qy	2563	AGTTCTCGGATCCGAGATGAGAAAGTGGATTGAGACATTCAGATGGCAATTGACCTGGGCG	2622
Db	714	AGTTCTCGGATCCGAGATGAGAAAGTGGATTGAGACATTCAGATGGCAATTGACCTGGGCG	655
Qy	2623	GAGAAAGACAGACAGCCCGCCCTGAGTTCTTGAGCAGACAGCCCGCTGACAAACAAGTCC	2683
Db	654	GAGAAAGACAGACAGCCCGCCCTGAGTTCTTGAGCAGACAGCCCGCTGACAAACAAGTCC	595
Qy	2683	CCTGATGAAGCCACCGCGGCTGATCCAGAGATCAGAGATGACCTGAGGGCTCTGGGCA	2742
Db	594	CCTGATGAAGCCACCGCGGCTGATCCAGAGATCAGAGATGACCTGAGGGCTCTGGGCA	535
Qy	2743	TGCGTGGAGCGCCACGAGCCCGGCAACCGGCGGCAACAATGATGACAGTGTGGGCAACCGC	2802
Db	534	TGCGTGGAGCGCCACGAGCCCGGCAACCGGCGGCAACAATGATGACAGTGTGGGCAACCGC	475
Qy	2803	AACACCAAGCGTCTCCATGATGGAATTTCAGATCCGAGTGAAGATCAGTTGTCTGGAAC	2863
Db	474	AACACCAAGCGTCTCCATGATGGAATTTCAGATCCGAGTGAAGATCAGTTGTCTGGAAC	415
Qy	2863	CTGCTGAGGAAATTCAAAAACAGCAACGGGATGGCAAGAGCTGTGGTGTGTTCAAAAC	2922
Db	414	CTGCTGAGGAAATTCAAAAACAGCAACGGGATGGCAAGAGCTGTGGTGTGTTCAAAAC	355
Qy	2923	TTCTGCGCTGTTCTTCTCAAAATTCACACGAGCAATATCCCGTTGCAGGCTGCTCTG	2982
Db	354	TTCTGCGCTGTTCTTCTCAAAATTCACACGAGCAATATCCCGTTGCAGGCTGCTCTG	295
Qy	2983	CTCGGCTACTCGCTCACACCATCCCTCTGATGCCGAAATTCAGAAAGAATCAGCTGTTC	3042
Db	294	CTCGGCTACTCGCTCACACCATCCCTCTGATGCCGAAATTCAGAAAGAATCAGCTGTTC	235
Qy	3043	AAGCTGCACTTCAAAGTCCACAGTCTACTTACAGGCGGAAAGGAGTACACGTTGCA	3102
Db	234	AAGCTGCACTTCAAAGTCCACAGTCTACTTACAGGCGGAAAGGAGTACACGTTGCA	175
Qy	3103	AGGTGAGATGGAAGATGCCGAGAGCCACAGAGCTCTGCTGGCCACCCCAAGTGTGAGC	3162
Db	174	AGGTGAGATGGAAGATGCCGAGAGCCACAGAGCTCTGCTGGCCACCCCAAGTGTGAGT	115
Qy	3163	CACAAAGAGTCTCTGTTGATTTGATGAGCGGAGACACATCGTTT-CCGAGATGGAGCTTT	3221
Db	114	CACAAAGAGTCTCTGTTGATTTGATGAGCGGAGACACATCGTTTCCGAGATGGAGCTTT	55
Qy	3222	CTTGGAAAGCGTTTCTTTCTTCTGTATTAAGAAGCTGTGTAATTAACACC	3275
Db	54	CCTGGAAAGCGTTTCTTTCTTCTGTATTAAGAAGCTGTGTAATTAACACC	1
RESULT 11			
BQ881649			
LOCUS			
DEFINITION			
BQ881649			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
1 (bases 1 to 912)			
BO881649 912 bp mRNA linear EST 16-AUG-2002			
AGNCOCURT 8728249 NIH_MGC_47 Homo sapiens cDNA clone IMAGE:5339275			
5', mRNA Sequence.			
BO881649			
BO881649.1 GI:22273657			
EST.			
Homo sapiens (human)			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;			
Homidae; Homo.			
(bases 1 to 912)			

FEATURES	source
<p> AUTHORS NIH-MGC http://mgc.nci.nih.gov/. TITLE National Institutes of Health, Mammalian Gene Collection (MGC) JOURNAL Unpublished (1999) CONTACT Robert Strausberg, Ph.D. EMAIL cgsabbs-remail.nih.gov TISSUE Tissue Procurement: ATCC CDNA Library Preparation: Rubin Laboratory CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA Sequencing by: Agencourt Bioscience Corporation DNA Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov PLATE: L1CM253 row: n column: 12 High quality sequence stop: 684. Location/Qualifiers 1. .912 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:6139275" /libase_type="neuroblastoma, cell line" /lab_host="DH10B (phage-resistant)" /clone_lib="NIH MGC 47" /notes="Organ: brain; Vector: pOT8; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH MGC Library." </p>	
<p> ORIGIN Query Match 25.7%; Score 883.4; DB 3; Length 912; Best Local Similarity 98.7%; Pred. No. 2.7e-225; Matches 901; Conservative 0; Mismatches 11; Indels 1; Gaps 1 </p>	
<p> 592 AAAAAATTAATCATATCACTACAGCAGAGCACTAGAGGACAAATCGTGAATTTCCACCAT 1 AAAAAATTAATCATATCACTACAGCAGAGCACTAGAGGACAAATCGTGAATTTCCACCAT 652 AACCAATTTGAGCAAAACACAGCAGAAATCAGATTTCCAGCTCTAGAGATTTGCCGCTCG 61 AACCAATTTGAGCAAAACACAGCAGAAATCAGATTTCCAGCTCTAGAGATTTGCCGCTCG 712 CTAAGAGATGATGGAATCCGGTTCCACCCGGCCAGAGCAGGAAAGGACAGAAATCAAT 121 CTAAGAGATGATGGAATCCGGTTCCACCCGGCCAGAGCAGGAAAGGACAGAAATCAAT 772 CTGGCCGTTGGCAACACGGGAATTTCTAGTGTTCAGGGTTTCATTAAGATCAATGCTCTC 181 CTGGCCGTTGGCAACACGGGAATTTCTAGTGTTCAGGGTTTCATTAAGATCAATGCTCTC 832 AACTGGGCGAAGGTGCGGAAGCTGAGCTTCAAGAGGAAGCGCTTTCTATCAAGCTCGG 241 AACTGGGCGAAGGTGCGGAAGCTGAGCTTCAAGAGGAAGCGCTTTCTATCAAGCTCGG 892 CCAATATGCAATATAGGTGATACAGAGATACCTTGGAAATTCCTGATAGGCCAGTCCGGATTC 301 CCAATATGCAATATAGGTGATACAGAGATACCTTGGAAATTCCTGATAGGCCAGTCCGGATTC 952 TGCAAGTCTCTTCTGGAATAATCTGTGTGGAATCATGATGCTCTTCTTGAAGCTTTTGAAGAG 361 TGCAAGTCTCTTCTGGAATAATCTGTGTGGAATCATGATGCTCTTCTTGAAGCTTTTGAAGAG 1012 CCAAAACCAAGCCCAAGCCCGTCTCTTTAGCCGGGGGTTCATCATTTCCGTTCAATGCT 421 CCAAAACCAAGCCCAAGCCCGTCTCTTTAGCCGGGGGTTCATCATTTCCGTTCAATGCT 1072 CGAGCTCAGAAAGCAGGTCTTCGATATAGTAAAGAGAGACATTAAGAGTGCAGATT 481 CGAGCTCAGAAAGCAGGTCTTCGATATAGTAAAGAGAGACATTAAGAGTGCAGATT </p>	<p> 11311 540 </p>

QY 1132 GAAAGAACACAGCAAGATTCATTCTATCCGAGCCTTGCTTCAACAGCCTACAGACTG 1191
| | | | |
DB 541 GAAAGAGACACAGCAAGATTCATTCTATCCGAGCCTTGCTTCAACAGCCTACAGACTG 600
| | | | |
QY 1192 AATTCGGAAGTCTGAGCAGTCTCAGCAGACACAGCCTTACATTGGAGAAGTGCC 1251
| | | | |
DB 601 AATTCGGAAGTCTGAGCAGTCTCAGCAGACACAGCCTTACATTGGAGAAGTGCC 660
| | | | |
QY 1252 GAATCTCCAGGGGGCCAGAGTCTCCGCGCAGAGAAAGAAACCGAAGGTTCCCGCGGGAG 1311
| | | | |
DB 661 GAATCTCCAGGGGGCCAGAGTCTCCGCGCAGAGAAAGAAACCGAAGGTTCCCGCGGGAG 720
| | | | |
QY 1312 CCGGGGTCGACCCGAGCCTTGCGCGCAGAGAAAGCCCGGGTTACAAGAGGCGAGC 1371
| | | | |
DB 721 CCGGGGTCGACCCGAGCCTTGCGCGCAGAGAAAGCCCGGGTTACAAGAGGCGAGC 780
| | | | |
QY 1372 GGAGCCGCTCGGCGCCACGAGAAAGAGAGAGTCTTTAAGATAGAACCCAGCAG 1431
| | | | |
DB 781 GGAGCCGCTCGGCGCCACGAGAAAGAGAGAGTCTTTAAGATAGAACCCAGCAG 840
| | | | |
QY 1432 AATTAACCTCAGCCCGCGAGCCAGACAGAGCTCCCTGACCTGGACCTTCC 1491
| | | | |
DB 841 AATTAACCTCAGGCCCCGAGCCAGACAGAGCTCCCTGACCTGGACCTTCC 899
| | | | |
QY 1492 GAGCTGTCTGTGA 1504
| | | | |
DB 900 GAGCTGTCTGTGA 912
| | | | |

RESULT 12
BU931960 904 bp mRNA linear EST 18-OCT-2002
LOCUS BU931960
DEFINITION AGENCOURT 10485886 NIH_MGC_109 Homo sapiens cDNA clone
IMAGE:6672254 5', mRNA sequence.
ACCESSION BU931960
VERSION BU931960.1 GI:24120779
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catherhini;
Homnidae; Homo.
1 (bases 1 to 904)
NIH-MGC http://mhc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Straubeberg, Ph.D.
Email: cga9ba-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: Agencourt Bioscience Corporation (LLNL)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CM2952 row: h column: 14
High quality sequence scop: 721.
Location/Qualifiers

FEATURES
Source
1..904

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6672254"
/tissue_type="teratocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: ovary; Vector: pOTB7; Site 1: EcoRI; Site 2:
XhoI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using Zap-cDNA synthesis kit (Stratagene) and

Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."

Query Match 25.4%; Score 873.2; DB 3; Length 904;
Best Local Similarity 98.6%; Pred. No. 1.5e-222;
Matches 891; Conservative 0; Mismatches 10; Indels 3; Gaps 1;

QY 1015 AAACCAAGCCCAACCCGTCCTTTAGCCGGGGGTCATCATTTCCGTTCAAGTGTGCG 1074
| | | | |
DB 1 AAACCAAGCCCAACCCGTCCTTTAGCCGGGGGTCATCATTTCCGTTCAAGTGTGCG 60
| | | | |
QY 1075 ACTGAAAGCAGGTTCTGACTATGTAAAGAGAGACATTAAGAGTGCAGTTTGA 1134
| | | | |
DB 61 ACTGAAAGCAGGTTCTGACTATGTAAAGAGAGACATTAAGAGTGCAGTTTGA 120
| | | | |
QY 1135 AGGAAGCAGCAGATTCATTTCTCCGAGCCTTGTTCACAGCCTTACAGATGTAAT 1194
| | | | |
DB 121 AGGAAGCAGCAGATTCATTTCTCCGAGCCTTGTTCACAGCCTTACAGATGTAAT 180
| | | | |
QY 1195 TCGAAGTGTGAGCAGTCTCAGCAGACACAGCCTTACATTTGAGAAAGTCCGAA 1254
| | | | |
DB 181 TCGAAGTGTGAGCAGTCTCAGCAGACACAGCCTTACATTTGAGAAAGTCCGAA 240
| | | | |
QY 1255 TCTCCAGGGGGCAGAGCTGCGCGAGAAAGAAACCGAAAGTTTCGCGGGAGCCG 1314
| | | | |
DB 241 TCTCCAGGGGGCAGAGCTGCGCGAGAAAGAAACCGAAAGTTTCGCGGGAGCCG 300
| | | | |
QY 1315 GGGTGGACCCGAGCGCTTGCGCCGAGAGAAAGCCCGGGGGTTAAACAGCGGAGCGA 1374
| | | | |
DB 301 GGGTGGACCCGAGCGCTTGCGCCGAGAGAAAGCCCGGGGGTTAAACAGCGGAGCGA 360
| | | | |
QY 1375 GCCGCTCGGCGCCCAAGAGAGAGAGAGTCTTTAAGATAGACCAAGCAGAGT 1434
| | | | |
DB 361 GCCGCTCGGCGCCCAAGAGAGAGAGAGTCTTTAAGATAGACCAAGCAGAGT 420
| | | | |
QY 1435 AAACCTCAGCCCCCGAGCCCAAGCAGAGCTCCCTGACCTGCAGCTCTTCCAG 1494
| | | | |
DB 421 AAACCTCAGCCCCCGAGCCCAAGCAGAGCTCCCTGACCTGCAGCTCTTCCAG 480
| | | | |
QY 1495 CTGTCTTGAACTCGCAGGGGGAGTGGCCCTGCAACGTGACCTTCCCAACTG 1554
| | | | |
DB 481 CTGTCTTGAACTCGCAGGGGGAGTGGCCCTGCAACGTGACCTTCCCAACTG 540
| | | | |
QY 1555 AGCCCGCAGCAGCAGAGCAGCTCTCTCTGATCAGCCGCTGCTGAATGACAGGCTGC 1614
| | | | |
DB 541 AGCCCGCAGCAGCAGAGCAGCTCTCTCTGATCAGCCGCTGCTGAATGACAGGCTGC 600
| | | | |
QY 1615 CCGCGAGCAGCAGATGAGAGAGGGCCGAGAGAAAGATTTCCAACTGATTAAGCGTAC 1674
| | | | |
DB 601 CCGCGAGCAGCAGATGAGAGAGGGCCGAGAGAAAGATTTCCAACTGATTAAGCGTAC 660
| | | | |
QY 1675 TTCAATAGCTAAGAGAGTCTTACCAACGAGGAAATATCTGAAGAGCTGGAAGTTATC 1734
| | | | |
DB 661 TTCAATAGCTAAGAGAGTCTTACCAACGAGGAAATATCTGAAGAGCTGGAAGTTATC 720
| | | | |
QY 1735 ACTTGCTGTTTCAGAGCAGTGTGCAAGAGAGCGCATGCGCGAAGCACTGAAAAGT 1794
| | | | |
DB 721 ACTTGCTGTTTCAGAGCAGTGTGCAAGAGAGCGCATGCGCGAAGCACTGAAAAGT 780
| | | | |
QY 1795 CTGATATTTCCGAATTTTGAACCTTTGCACAAAATTTCTAATATTTTCTCAAGAAAT 1854
| | | | |
DB 781 CTGATATTTCCGAATTTTGAACCTTTGCACAAAATTTCTAATATTTTCTCAAGAAAT 840
| | | | |
QY 1855 GAGCAACAGCTTGCCCTGTTGGAGAGGCGCGCTCAAAATGCCCAATAGAGATTACCA 1911
| | | | |
DB 841 GAGCAACAGCTTGCCCTGTTGGAGAGGCGCGCTCAAAATGCCCAATAGAGATTACCA 900
| | | | |
QY 1912 AGAA 1915
| | | | |
DB 901 AGAA 904
| | | | |

RESULT 13	
LOCUS	BMS51766
DEFINITION	BMS51766 1026 bp mRNA linear EST 20-FEB-2002
ACCESSION	AGNCOURT_6575294 NIH_MGC_98 Homo sapiens cDNA clone IMAGE:5479350
VERSION	5'', mRNA sequence.
KEYWORDS	BMS51766 GI:18789071
SOURCE	EST.
ORGANISM	Homo sapiens (human)
	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
	Homnidae; Homo
REFERENCE	1 (bases 1 to 1026)
AUTHORS	NIH-MGC http://mgc.ncl.nih.gov/.
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL	Unpublished (1999)
COMMENT	Contact: Robert Strusberg, Ph.D. Email: cgapbs-remail.nih.gov Tissue Procurement: ATCC cDNA Library Preparation: Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNLI) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNLI at: http://image.llnl.gov Plate: L1CM2000 row: h column: 07 High quality sequence stop: 675.
FEATURES	location/Qualifiers
source	1..1026
	/organism="Homo sapiens"
	/mol_type="mRNA"
	/db_xref="taxon:9606"
	/clone="IMAGE:5479350"
	/tissue_type="astrocytoma grade IV, cell line"
	/lab_host="RDH10B (phage-resistant)"
	/clone_id="NH_MGC_98"
	/note="Organ: brain; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGGCAGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC library."
ORIGIN	
Query Match	25.1%; Score 863.8; DB 2; Length 1026;
Best Local Similarity	96.7%; Pred. No. 5.2e-220;
Matches 916, Conservative	0; Mismatches 22; Indels 9; Gaps 3;
OY	2045 GAGACTTGAGCTGAGAGAAGGTGTTACCTACCGCTCAACACCCTTCCTCGGGGCCAC 21040
Db	1 GAGACTTGAAGTGCAGAAAGGTGTATCTTACCGCTCAACACCCTTCCTCGGGGCCAC 60
OY	2105 TGCAACCGGCTCATGATCACTACAAGCAGTCTCTGAAGCGGCTGTGCAACACCAACCGCGCGA 2164
Db	61 TGCACCGGCTCATGATCACTACAAGCAGTCTCTGAAGCGGCTGTGCAACACCAACCGCGCGA 120
OY	2165 GCCACGCGGACTTCAGGAGACTGCCGAGCGCGCTTTGGCAGAGATCACGAGATGGTGGCAC 2224
Db	121 GCCACGCGGACTTCAGGAGACTGCCGAGCGCGCTTTGGCAGAGATCACGAGATGGTGGCAC 180
OY	2225 AGCTCCACGGTAGCATGTATCAAGATGGAGATTTCBAACCTGTGACGAACCTCAAGAAAG 2284
Db	181 AGCTCCACGGTAGCATGTATCAAGATGGAGAAATTTCAAGAGTGCACCACTCAAGAAAG 240
OY	2285 ATTGATTTGGCATTTGACATCTTTGTGTTCGGGAAGGAGTTCACTCCGTCTGGGCAACC 2344
Db	241 ATTGATTTGGCATTTGACATCTTTGTGTTCGGGAAGGAGTTCACTCCGTCTGGGCAACC 300
OY	2345 TCAGCAAGCTCTCGGGGAAGGGGCTCCAGCAGCGCATTTCTTCTGTTCACGACGTCC 2404

Db	301	TCAGCAAGCTCTGGGGGAAAGGGGGCTCCAGACAGCGCAGTATTCTTCTGTCAACGACGCTCC	360
QY	2405	TGCTATACACGAGCCCGGGGGCTGACGGCTTCCATTCAGTTTAAAGTCCACGCGGACGCTCC	2464
Db	361	TGCTATATACACGAGCCGGGGGGCTGACGGCTTCCATTCAGTTTAAAGTCCACGCGGACGCTCC	420
QY	2465	CGCTCTATGGCATACGATTGAGGAGAGCGAAGACGAGTGGGGGGTCCGCCACTGCGCTGA	2524
Db	421	CGCTCTATGGCATACGATTGAGGAGAGCGAAGACGAGTGGGGGGTCCGCCACTGCGCTGA	480
QY	2525	CCCTCCGGGGCCACGGGACGTCATTCATTCGTGGCCGGCAGTTCTCGATCCGAGATGAGAGA	2584
Db	481	CCCTCCGGGGCCACGGGACGTCATTCATTCGTGGCCGGCAGTTCTCGATCCGAGATGAGAGA	540
QY	2585	AGTGGGTTGAGGACATCCAGATGGCCATTGACTCTGGCGGAGAGACAGACGCCCGCCC	2644
Db	541	AGTGGGTTGAGGACATCCAGATGGCCATTGACTCTGGCGGAGAGACAGACGCCCGCCC	600
QY	2645	CTGAGTTCCTGGCCAGACGCCCTCCACAAAGTCCCTCGATGAAGACACCGCGGCTG	2704
Db	601	CTGAGTTCCTGGCCAGACGCCCTCCACAAAGTCCCTCGATGAAGACACCGCGGCTG	660
QY	2705	ACGAGAGTCAGAGGATGACTCTGAGCGCCTTCGCGCACATCGCTGAGAGCGCACGCCCGC	2764
Db	661	ACGAGAGTCAGAGGATGACTCTGAGCGCCTTCGCGCACATCGCTGAGAGCGCACGCCCGC	720
QY	2765	ACCGCGGCAACACATGGTGTGCTGTGCTGGACCGGCAACACGAGGCTTCCATGGTGG	2824
Db	721	ACCGCGGCAACACATGGTGTGCTGTGCTGGACCGGCAACACGAGGCTTCCATGGTGG	780
QY	2825	ACTTCAGCATGTGCGAGTGGAGAAATGATTTGTCTGGAAACCTGCTGAGGAAATTCAAAAACA	2884
Db	781	ACTTCAGCATGTGCGAGTGGAGAAATGATTTGTCTGGAAACCTGCTGAGGAAATTCAAAAACA	840
QY	2885	GCAAC-GGGTGGCAGAAAGCTGTGGGTGTGTTCAAACTTCTGCTG-----TTCTCT	2938
Db	841	GCAACGGGGTGGCAGAAAGCTGTGGGTGTGTTCAAAACTTCTGCTGCTTCTCT	900
QY	2939	ACAATATCACACGAGACAT---CATCCCTTTCGACGCTGCTCTG	2982
Db	901	ACAATATCACACGAGACATTCATTCCTCCCTTGCCACGCTGCTCTG	947

RESULT 14
AL566821/c
LOCUS
DEFINITION AL566821 Homo sapiens FETTL BRAIN Homo sapiens cDNA clone
ACCESSION CS0DF025Y01.3-PRIME, mRNA sequence.
VERSION AL566821
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
 Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Hominidae; Homo.
 li, W.B., Gruber, C., Jesssee, J. and Polayres, D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 2 rue Gaston Cremieux, CP 5706 - 91057 Evry cedex - FRANCE
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
 end enriched, double-strand cDNA was digested with Not I and cloned
 into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
 was not normalized. Library was constructed by life Technologies, a
 division of Invitrogen.
 This sequence belongs to sequence cluster 7229.f
 For more information about this cluster, see
 http://www.genoscope.cns.fr/cdnas?CS0DF025CE01NP1&c=7229.f.
 Location/Qualifiers

FEATURES

```

source
1. .930
/organism="Homo sapiens"
/mol_type="rRNA"
/db_xref="taxon:9606"
/clone="CS0DF025J01"
/issue_type="FETAL BRAIN"
/dev stage="fetal"
/clone_1b="Homo sapiens FETAL BRAIN"
/notes="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo (df) primer. Five prime end
enriched, double-strand cDNA was digested with NotI and
cloned into the NotI and EcoRI sites of the pCMVSPORT 6
vector. Library was not normalized."

ORIGIN
Query Match      25.0%; Score 862; DB 1; Length 930;
Best Local Similarity 98.2%; Pred. No. 1.5e-219;
Matches 915; Conservative 0; Mismatches 10; Indels 7; Gaps 4;

2366 GGCTCCAGCAGCGCATGTTCTTCTGTTCAACGACGCTGCTATACAGACCGGGGGC 2425
2370 GGCTCCAGCAGCGCATGTTCTTCTGTTCAACGACGCTGCTATACAGACCGGGGGC 871
2426 TGAACGCGCTCCATATCAAGTTAAAGTCACAGGAGCTCCCGCTTATGCGATGACGATTG 2485
870 TGAACGCGCTCCATATCAAGTTAAAGTCACAGGAGCTCCCGCTTATGCGATGACGATTG 811
2486 AGGAGGCGGAAGACGAGTGGGGGGTCCCACTGCTGACCTCCGGGGGCGACGGGCACT 2545
810 AGGAGGCGGAAGACGAGTGGGGGGTCCCACTGCTGACCTCCGGGGGCGACGGGCACT 752
2546 CCATCATCGTGGCGCGCATGTTCTGCTCGAGATGAGAAAGTGAGTGAAGCATCCAGA 2605
751 CCATCATCGTGGCGCGCATGTTCTGCTCGAGATGAGAAAGTGAGTGAAGCATCCAGA 692
2606 TGGCCATTGACCTGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2665
691 TGGCCATTGACCTGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 632
2666 CCCCTGACAAACAGTCCCTGATGAGAGCAACGCGGCTGACAGAGATGAGAGAGAG 2725
631 CCCCTGACAAACAGTCCCTGATGAGAGCAACGCGGCTGACAGAGATGAGAGAGAG 572
2726 TGAAGCGCTTCGCGACATCGCTGAGAGCGGCGCGGCGGCGGCGGCGGCGGCGG 2785
571 TGAAGCGCTTCGCGACATCGCTGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGG 513
2786 ACGTGTGCTGGACCGCAACACGAGCTTCCATGAGTGAAGTTCA--GCATCGAGTGG 2842
512 ACGTGTGCTGGACCGCAACACGAGCTTCCATGAGTGAAGTTCAAGCGCGCGGCGG 453
2843 AGAATAGTGTCTGGAACCTGCTGAGAAATTTAAACAGCAAGGAGGAGGAGAGAG 2902
452 AGAATAGTGTCTGGAACCTGCTGAGAAATTTAAACAGCAAGGAGGAGGAGAGAG 393
2903 TGTG--GATGATGTTCAAACTTCTGCTGCTTCTTCAAAATCAACACAGAGACATCA 2960
392 TGTGAGGAGGAGTTCACAACTTCTGCTGCTTCTTCAAAATCAACACAGAGACATCA 333
2961 TCCCTTTCGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3020
332 TCCCTTTCGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 273
3021 CATTCAGAAAGACTAGTGTTCAGAGTGAAGTTCAGAGTTCAGAGTTCAGAGTTCAG 3080
272 CATTCAGAAAGACTAGTGTTCAGAGTGAAGTTCAGAGTTCAGAGTTCAGAGTTCAG 213
3081 GGAAGAGGAGTACACCTTCGAAAGGTGATGAGAGTATCCGAGTTCAGAGTTCAG 3140
212 GGAAGAGGAGTACACCTTCGAAAGGTGATGAGAGTATCCGAGTTCAGAGTTCAG 153
3141 CTGCGAGACCCACGATGTTAGCCAAAGAGTCTTTGTGATGATGAGCGGAGCACT 3200

```

```

Db      152 CTGCGGACCCGACGATGTTAGTCAACAAAGTCTCTGTGATGATGAGCGGAGACACT 93
QY      3201 CGTTCCGAGAGGCTGCTTCTGGAAGAGCTTCTTCTGATTAATGAGAGCT 3260
Db      92 CGTTCCGAGAGGCTGCTTCTGGAAGAGCTTCTTCTGATTAATGAGAGCT 33
QY      3261 GGTAAATTAACACCTGCTGGAATCAAAA 3292
Db      32 GGTAAATTAACACCTGCTGGAATCAAAA 1

RESULT 15
AY415515
LOCUS
DEFINITION
Homo sapiens PARP2 gene, VIRUAL TRANSCRIPT, partial sequence,
AY415515
VERSION
AY415515.1 GI:39771474
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Mammalia; Eutheria; Barchontoglires; Primates; Catarrhini;
Homnidae; Homo.
1 (bases 1 to 3165)
Clark,A.G., Gianwakt,S., Nielson,R., Thomas,P., Kejarival,A.,
Todd,M.A., Tanendau,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Shtinsky,J.J.,
Adams,M.D. and Cargill,M.
Interfering nonneutral evolution from human-chimp-mouse orthologous
gene trios
Science 302 (5652), 1960-1963 (2003)
JOURNAL
PUBMED
14671302
REFERENCE
AUTHORS
Clark,A.G., Gianwakt,S., Nielson,R., Thomas,P., Kejarival,A.,
Todd,M.A., Tanendau,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Shtinsky,J.J.,
Adams,M.D. and Cargill,M.
Direct Submission
TITLE
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT
This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source
1. .3165
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1. .>3165
/gene="PARP2"
/locus_tag="HCM5583"

ORIGIN
Query Match      24.5%; Score 842.4; DB 14; Length 3165;
Best Local Similarity 51.9%; Pred. No. 4.1e-214;
Matches 1606; Conservative 0; Mismatches 144; Indels 45; Gaps 7;

69 GCCGACCCGAGATCAACGACTGCGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 128
33 GCGAGCTGAGGAGATGCGCTTGGGTGCCAGACCCCTGTGGAGTTAGACCCCTGAGCC 92
129 TGAAGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 188
93 TGGGAGAGCTCTCTGCGGAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 152
188 GAGTACACCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 248
153 GAGACACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 212
249 TGCAGTTGCAACCACTCAACCTGCTGAGAGAGTACTATTTGGCTGAGATTTCTGA 308
213 ACAAGTGGAGAGCGTTTAACTGTAGATGATGATGATGATGATGATGATGATGATG 272

```

Oy	309	TCACAAAAAGATCAGCGTGTGGCTGGATCTCTAAACCATTGTGGAAACAGATTGAAG	368
Db	273	TACTCAGTCTTACTGGAAATTTGGCTTGAACCTATGAAACCCATCAATTAGGCAAAATCGAAG	332
Oy	369	GCCAAAGACAGTGTGTTGAATTTAACTTGTGGTGAATTTCTTCCGGCTGACACACAACT	428
Db	333	GCCAAAGAAATGTGGTCTTCCGCTTACGTGTAAATTTTTCACCTTGATCTGGTCAAGT	392
Oy	429	CCAAAGAACTCACAAAGTACCTTTCGCGCTGACAGGTGAAGACGAGACTTGGCTCAAG	488
Db	393	ACAAAGAAATATACAGATTACTTGTGGCTTGGACACTTAAGAGAGACCTGTGGAAAG	452
Oy	489	CAGGTTCAGCTGTAAATGACACGAGCGGAGCTCTTGTATTCACACATTTGTCAATCTGA	548
Db	453	GCCTTTCAGCTGTGTGCTGACACCCACAGCGGCGCTTCTCAAGTCCCTTCGCACTGCACTCGGA	512
Oy	549	GATTTGGGATTTTGAATGAAGCCTTGGACAGAGAGACCTTAGCAAAAATTAATATATCC	608
Db	513	AATTAAGAGATTAAGATGAACCGCTGACCGAAGACCTCAAAAGTGAACGATATTTCC	572
Oy	609	TCAGCAAGACGACCTAGAGACCAAAATCGTGGAAATTTGACATAACCAATTTGACAAC	668
Db	573	TGGCCAGACGACCTCCCTTGAAGAAATTAAGATTTCCATCAGAAGACGTGGGCAAC	632
Oy	669	ACCAGCAGAAATCAGATTTTCCAGCTCTAGAGATTGCGCGGTGAGATGTATGAAT	728
Db	633	ACCTGCTAGTCGGAATTTCCAGGTGCTCGAAATTTCTCGAAAGTTGAAATGTACGGCAT	692
Oy	729	CCGCTTGAACCCGGCCAAAGACAGAGGAAGGACAGAAATCAATCTGGCGCTTGGCAAC	788
Db	693	CAGATTTACATGCTGCTTCTGACAGGGAAGAAACCAAGATTCACATGCGAGTTTCCACAT	752
Oy	789	GGAATTTCAATGTTTCAGGGTTTCACTAAGATCAATGCTTCACTGGGCAAGTGC	848
Db	753	GGGTGTACTCGTGTTCAGANN	812
Oy	849	GAACTGAGCTTCAAGAGAAACGCTTCTCATCAAGCTCCGGCCAGATGCCAATATGTC	908
Db	813	NN	872
Oy	909	GTACCAGGATACCTTGGAAATTCCTGATGGCCAGTGGGATTTCTGCAAGTCTTCTGAA	968
Db	873	TTACCAGGACATTTGAATTTTGTGGTGAATGAATGAATGAATGAATGAATCTTGGAA	932
Oy	969	AATCTGTGTAACATCATGCTTCTTGAATTTTGAAGGCCCAACCAAGCCCAA	1028
Db	933	GATTTGTGTGAAGTATCAACCTTTTGAACCTTTTGAACCAACTTAAGCCAAAGCAAA	992
Oy	1029	GCCCGTCTCTTTAGCCCGGGGATCATTTTCGTTCAGTGTCGACTCAAGACAGT	1088
Db	993	AGCGGTCTTCTTACGCGGGGCTCTCTCTTCAATACAGANNNNNNNNNNNNNNNNNNNN	1052
Oy	1089	TCTGACATATTTTAAAGAGAGACATTAAGAGTGCAGTTTGAAAGAAACACAGCA	1148
Db	1053	NN	1112
Oy	1149	GATTCATTTCTACCGGAGCTTGTCTCAAGGCTTACAGATTCGAATTCGAAGTGTGGA	1208
Db	1113	GACCCACACGCTC-----GTTGACCTCTGACTGACACTTACAA	1154
Oy	1209	GCACTCTCAGACGAGCACAGCTTACATTTTGAAGAGTGCAGATCTCCAGGGGGGCA	1268
Db	1155	ACAGAGCATTCATTTCCCGGAGGATTTGAGGACTCTGCTCCCATTTTGAAGGAATGC	1214
Oy	1269	GAGCTGCCGCGAGGAAGAAACGAAAGTTTCCGCGGGGAGCGCGGGGTGCACCCGAG	1328
Db	1215	CTTTTATCTCGTCTCTCTCCCTCACCTGTGCTCCCTCTGCGCAGAGTTTAAAGACAG	1274
Oy	1329	CCCTGCGCGAGAGGAAGCCCCCGGGGTAACAAGAGCGGACGGAAGCGCTGGGCGC	1388
Db	1275	CAGCAGCTCCTTCAACAGATTCCTCCAGGTTTCTTAAGTCAAGAGTCCAGTGTGAAGAGCG	1334
Oy	1389	CACGAGAGAGAGAGAG--TCTGTTAAGATAGACCCAGACAGATTAACCTACGCC	1445

Db	1335	CAGTGAAGCAGTGGCTGAGGCCCCGACACACCATCGGCTCACCCCTCGGGCCCCCGC	1394
Qy	1446	CCCGGAGCCAAACACACAGGCTCCCTGATCGGACGTCCCACTTTCCGAGCTGTGTGAA	1505
Db	1395	ACTCCAGCTGGTCCAGGCTTTTCCAGGAAGATCTCAGACCTTCTCCCTCCAGCCGAA	1454
Qy	1506	CTCGCAGGGGGAGTGGCCCCCTTCGCAACGTGACTTGTCTCCGACCTGAGCCCCGAC	1565
Db	1455	GAGCCCCCTGAGCTGAGCCCTGCAATTCAAGTGGCTT-----TGGGCCAGCTGA	1505
Qy	1566	CAGCAGGCTCTCCCTTGAATCAGCCCCGTCTGTAATGACCAAGCTCTGCCCCGGAACGA	1625
Db	1506	ACAGGGCTCATCCCACTCCTGAGCCCTGTCTCACTGATGCTGGCGGAGCCGGAGTGA	1565
Qy	1626	CGATGAGATGAGGGGCGGAGGAGATTTCCCACTGATTAAGCGTACTTCACTAGCTAA	1685
Db	1566	CTGCGAG--GAGCCAGACACAGGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	1622
Qy	1686	GGAAGTGTCTACACCGAGCGAACATATCTGAAGATCTCGAATTATCACTTCGTGTT	1745
Db	1623	NNNTGGTT	1682
Qy	1746	TCAGACACATGAGCAAAAGAGAGCCCATGCCGGAAGCATGAAAAAGTCTCATTTCCC	1805
Db	1683	CCGACGCGCAGTGGTGAAGAGAGACCCATGCTGCGACTGTGATACGCTGCTCTCTC	1742
Qy	1806	GAAATTTGAACTTTTGCACAAAATTCATACTAATTTTTCAGAGAAATTGAGACGACT	1865
Db	1743	CACATCGATCCCATCTATGAGTTCCACAGAGGCTTCTGCGGAGGTGAGCAGAGCT	1802
Qy	1866	TCGCTGTGGGAAGGCGCTCAATATGCCCAATCAGAG--ATTACCAAGAAATCGCGCA	1922
Db	1803	GGCATCTGGGAAGGGCCCTCCAAAGCCACACAAAGGCAATCATAGAAATCGGGA	1862
Qy	1923	TGTCAATGCTGAAGAACATTCAAGGCGATGAGCAGCTGGCGGCTCACCTGTGAGACAG	1982
Db	1863	CATCTGCTCAGGAACATGCGCAAGTTAAAGGAGTTTACAGCTACTTCCAAAGACATGA	1922
Qy	1983	CGAGGCTTTGAGGCGCTGGAGATGGAATGCAAGACTCCGCGGCTGGAGAACTTCTG	2042
Db	1923	CGAGGTCTTAAGAACTGGAAGAGCTACAAACGCTGTGAAGATTTGAGGACAGTGA	1982
Qy	2043	CAGAGACTTTGAGCTGAGAAAGTGTGTTACTTACGCTCAACACTTCTCTGCGGCC	2102
Db	1983	CAGAGATTTGAGCTGAGAAAGTGTGTTACTTCTCTCAACACTTCTCTGTTGAAGCC	2042
Qy	2103	ACTGACCGGCTCATGCACTACAGAGAGGTCTTGAAGCGGCTGTGAAACACCAACCGGC	2162
Db	2043	CATTCAGGCGCTGCTGCACTACCGCTGTGCTGCGCCGCTTATGCGACATTAACAGCC	2102
Qy	2163	GAGCCAGCCGCACTTCAGGGAATGCGGAGCCGCTTTGAGAGAGATCAAGAGATGTGTGC	2222
Db	2103	CGGGACCATGACTGACGTGACCTATGACGCCCTGAAAGCCATACAGAGGTGACAC	2162
Qy	2223	ACAGCTCACGCGTATGATGATCAAGATGAGAAATTTCCAGAAAGCTGACAGAACTCAAGAA	2282
Db	2163	CACACTACAGCACTTCTCATCGGCTGGAAGAACTGCAAGAACTTAAAGAGCTGACAGC	2222
Qy	2283	AGATTGATTTGGCATTGAACATCTTGTGTTCCTGGAGAGGAGTTATCCGTCTGGCAG	2342
Db	2223	GGAAGCTGGTGGCATTAAGAAACCTCATTTCTCTCGCAGGAGTTATCCGTGAGGGCTG	2282
Qy	2343	CCTCAGCAAGCTCTGGGGGAGGGGCTCCAGCAGCGCATGTTCTTCTGTTCACAGCT	2402
Db	2283	CCTTCACAGCTGACCAAGAAAGGGCTGACAGCAGAGAGTGTCTGTCTTCACATAT	2342
Qy	2403	CTGCTATATACAGAGCCGGGGGCTGACGGCTCCCATATGAGTTTAAAGTCAACGGGAGCT	2462
Db	2343	GTTGCTGTATACAGCAAGAAAGAGTTGACGGAGCAAGCCATTCGGAATCGGGGCTTCT	2402
Qy	2463	CCCGCTTATGCGATGAGATTTGAAGAGACGAAGACGAGTGGGGGTGCCCATGCTCT	2522

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using bw model

Run on: May 22, 2006, 00:49:20 ; Search time 595.56 Seconds
(without alignments)
10813.957 Million cell updates/sec

Title: US-09-555-342B-1

Perfect score: 3442
Sequence: 1 CGCCGAGCGCGCGCGCGCTG.....AGTAAACATGTCATC 3442

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1403666 seqs, 935554401 residues

Total number of hits satisfying chosen parameters: 2807332

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Issued Patents NA:*

- 1: /EMC Celerra_SIDS3/ptodata/2/ina/1/COMB.seq:*
- 2: /EMC Celerra_SIDS3/ptodata/2/ina/5/COMB.seq:*
- 3: /EMC Celerra_SIDS3/ptodata/2/ina/6A/COMB.seq:*
- 4: /EMC Celerra_SIDS3/ptodata/2/ina/6B/COMB.seq:*
- 5: /EMC Celerra_SIDS3/ptodata/2/ina/7/COMB.seq:*
- 6: /EMC Celerra_SIDS3/ptodata/2/ina/H/COMB.seq:*
- 7: /EMC Celerra_SIDS3/ptodata/2/ina/PC/US.COMB.seq:*
- 8: /EMC Celerra_SIDS3/ptodata/2/ina/PP.COMB.seq:*
- 9: /EMC Celerra_SIDS3/ptodata/2/ina/RE.COMB.seq:*
- 10: /EMC Celerra_SIDS3/ptodata/2/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	238	6.9	310	3	US-09-621-976-8193	Sequence 8193, App
2	192.4	5.6	2853	3	US-09-949-016-1933	Sequence 1933, App
3	180.4	5.2	6263	3	US-09-664-958-5	Sequence 5, Appl
4	180.4	5.2	6268	3	US-09-566-921-57	Sequence 57, Appl
5	169.6	4.9	4292	3	US-09-949-016-1488	Sequence 1488, App
6	169.6	4.9	4336	3	US-09-949-016-208	Sequence 208, App
7	165.8	4.8	576	3	US-09-270-767-1995	Sequence 1995, App
8	165.8	4.8	576	3	US-09-270-767-1777	Sequence 1777, A
9	144	4.2	3984	3	US-09-848-294-1	Sequence 1, Appl
10	143.2	4.2	2872	3	US-09-906-779-3	Sequence 3, Appl
11	142.4	4.1	2156	3	US-09-949-016-3593	Sequence 3593, App
12	133.6	3.9	3398	3	US-09-799-451-201	Sequence 201, App
13	106.4	3.1	449	3	US-09-270-767-1542	Sequence 1542, App
14	106.4	3.1	449	3	US-09-270-767-16824	Sequence 16824, A
15	105.2	3.1	184	3	US-09-621-976-15728	Sequence 15728, A
16	97.8	2.8	17098	3	US-09-921-976-15864	Sequence 15864, A
17	95	2.8	2493	3	US-10-104-047-41	Sequence 41, Appl
18	90.6	2.6	573	3	US-09-270-767-165	Sequence 165, Appl
19	90.6	2.6	573	3	US-09-270-767-15447	Sequence 15447, A
20	65.2	1.9	3438	3	US-10-164-595-29	Sequence 29, Appl
21	63.8	1.9	4272	3	US-09-949-016-5747	Sequence 5747, App
22	63.8	1.9	1626	3	US-09-620-312D-1033	Sequence 1033, App
23	63.8	1.9	1699	5	US-09-913-171A-26	Sequence 26, Appl

24	62.8	1.8	2730	3	US-09-774-528-175	Sequence 175, App
25	62.8	1.8	2730	4	US-10-120-988-175	Sequence 175, App
26	62.8	1.8	3503	4	US-10-094-749-560	Sequence 560, App
27	59.6	1.7	722	3	US-09-270-767-1861	Sequence 1861, App
28	59.6	1.7	722	3	US-09-270-767-17143	Sequence 17143, A
29	59.4	1.7	222	3	US-09-016-434-45	Sequence 45, Appl
30	56.8	1.7	4080	2	US-08-446-345-35	Sequence 35, Appl
31	56.8	1.7	4081	3	US-09-949-016-1292	Sequence 1292, App
32	53.8	1.6	7218	2	US-08-232-463-14	Sequence 14, Appl
33	50.2	1.5	3044	3	US-09-814-915A-84	Sequence 84, Appl
34	50.2	1.5	3044	4	US-09-880-107-3718	Sequence 3718, App
35	48.8	1.4	1308	3	US-09-252-991A-11936	Sequence 11936, A
36	48.8	1.4	1335	3	US-09-252-991A-12073	Sequence 12073, A
37	48.2	1.4	1816	3	US-09-620-312D-439	Sequence 439, App
38	48.2	1.4	3415	3	US-09-902-540-5107	Sequence 5107, App
39	48.2	1.4	28058	3	US-09-902-540-1252	Sequence 1252, App
40	47.4	1.4	2400	3	US-08-930-001-1	Sequence 1, Appl
41	47.4	1.4	2400	3	US-09-091-885-1	Sequence 1, Appl
42	47	1.4	801	3	US-09-902-540-8879	Sequence 8879, App
43	47	1.4	12028	3	US-09-902-540-1071	Sequence 1071, App
44	47	1.4	54711	3	US-09-949-016-17489	Sequence 17489, A
45	46	1.3	235452	3	US-09-949-016-13675	Sequence 13675, A

ALIGNMENTS

RESULT 1
US-09-621-976-8193
Sequence 8193, Application US/09621976
Patent No. 6639063
GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, J.B.
APPLICANT: Joberet, S.
APPLICANT: Giordano, J.Y.
TITLE OF INVENTION: ESTs and Encoded Human Proteins.
FILE REFERENCE: GENSET.054PR2
CURRENT APPLICATION NUMBER: US/09/621, 976
CURRENT FILING DATE: 2000-07-21
NUMBER OF SEQ ID NOS: 19335
SOFTWARE: Patent.pm
SEQ ID NO 8193
LENGTH: 310
TYPE: DNA
ORGANISM: Homo sapiens
US-09-621-976-8193

Query Match
Best Local Similarity 95.3%; Pred. No. 1.6e-53;
Matches 244; Conservative 10; Mismatches 1; Indels 1; Gaps 1;

QY	1	CGCCGAGCGCGCGCGCT-GTGAGATATTTCTTAAGCCGCTTTCATCATGGAGAAT	59
DB	55	CGCCGAGCGCGCGCGCTGTGAGATATTTCTTAAGCCGCTTTCATCATGGAGAAT	114
QY	60	AAGAGAGAGCGAGCCGATCCGATCCGATCGAGGCGCGGAAATTCGGGATCAGTAC	119
DB	115	AAGAGAGAGCGAGCCGATCCGATCCGATCGAGGCGCGGAAATTCGGGATCAGTAC	174
QY	120	CTTGGAACTGTGACAGAAAGCCGCCCAACCTTCAGAAACTGTCTCATCAAAAT	179
DB	175	CTTGGAACTGTGACAGAAAGCCGCCCAACCTTCAGAAACTGTCTCATCAAAAT	234
QY	180	CCAGATCTGTGATGACCCAGAGAGCATTTAAGTTCACAAAGAGCTCTGGGAAGT	239
DB	235	CCATATGCTGTGATGACCCAGAGAGCATTTAAGTTCACAAAGAGCTCTGGGAAGT	294
QY	240	GCTGCTGATGAGATT 255	
DB	295	GCTGCTGATGAGATK 310	

RESULT 2

US-09-949-016-1933
 ; Sequence 1933, Application US/09949016
 ; Patent No. 6812339
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMOPLHSMS IN KNOWN GENES ASSOCIATED
 ; FILE REFERENCE: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CL001307
 ; CURRENT APPLICATION NUMBER: US/09/949,016
 ; CURRENT FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: 60/241,755
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/237,768
 ; PRIOR FILING DATE: 2000-10-03
 ; PRIOR APPLICATION NUMBER: 60/231,498
 ; PRIOR FILING DATE: 2000-09-08
 ; NUMBER OF SEQ ID NOS: 207012
 ; SOFTWARE: FASTSEQ for Windows Version 4.0
 ; SEQ ID NO 1933
 ; LENGTH: 2853
 ; TYPE: DNA
 ; ORGANISM: Human
 ; US-09-949-016-1933

Query Match 5.6%; Score 192.4; DB 3; Length 2853;
 Best Local Similarity 52.8%; Pred. No. 7,8e-41;
 Matches 490; Conservative 0; Mismatches 426; Indels 12; Gaps 3;

165 CGTGTCCATCAAAATCCAGATGCTGGATGACACCCAGGAGGCAATTGGAAGTCCCAAAAG 224
 784 CATGACATGCAAGGTTCTTGTGTGATGACACAGTTTATGATGTGTGAGAGAAACA 843
 225 AGCTCTGGGAGGAGTCTGCTGGATGCAACCACTCAACCTGCGGAAAGTGA 284
 844 TGCTAAGGACAAAGATTGCTTAAAGAGATGATGAGCATCTCAATCTTTGGAAGAGA 903
 285 CTATTTTGGCTGAGTTTCTGATCACAAAAAGATCAGGTGTGGCTGGATCTCTAAA 344
 904 CTATTTTGGCTGAGTTTGGATTAACCACTCTAAGACATGGCTGATTCGCCCA 963
 345 ACCATTGTGAACGATTTAGAAAGGCAAGCAAGCTGTGTGTAAGTTGTGTGTAAGT 404
 964 AAAAAATAAAAAGAGGTCTGTGG--TGTCCCTTGGAAATTTTACATTTAATGTAAGTT 1020
 405 CTTTCGCTGACCAACACCACTCCAGAGAAATCAACAGTAAGTGTGCGCTGCA 464
 1021 TTATCACTGACCAAGCAAGTAACAGAAACATTAACAAATATTTATGTCTTCA 1080
 465 GGTGAAGCAGACTTGGCTCAAGGAGTTGACGTGTAAATGACACCAAGCGACGCTCTT 524
 1081 GCTTGGGAGGACATATGTTGACAGAGCTGTGCCCTGTCTTGGCAACCTTAGCATTAAT 1140
 525 GATTTCACACATTTGCAATCTGAGATTTGGGAGATTTTGAAGCTTGGACAGAGAGA 584
 1141 AGGTCTTACACATCTGACGTGAACTGAGAGACTAGCAAGCAAGCTTCATGCGGTGA 1200
 585 CTTAGCAAAAATTAATACATAC-----TCAGCAAGACGACTAGAGACAAATCGT 638
 1201 TTATGTAGTATTTTAACTGGCCCCGAATCAAGCAAGAACTTGAAGAGAAAGTCAAT 1260
 639 GGAATTTCAACATTAACACATTTGACAAACACACAGAGAAATGATTAATTCAGCTCTTGA 698
 1261 GGAATGCAATATGATACAGGTTCATGACTCAGCTCAGCTGACTGACTTGGATTTCTTGA 1320
 699 GATTGCCCGTGGCTAGAGATGATGAGATCCGGTTGACCCGGGCAAGAGAGAGAGAGG 758
 1321 GAATGCCAAAAGTGTCTATATATGAGATTAATCTTCAAAAGCAAGAACTTGAAGG 1380
 759 CACGAAGATCAATCTGCGCTTGGCAACACGAGAAATTTAGTGTTCAGGGTTTCACTAA 818
 1381 AGTAAATATCAATCTGAGTGTCTGCTTAGTGCCCTTCTGCTTTACAAAGATTAAGCTAG 1440
 819 GATCAATGCTTCAACTGGGCAAGGTGCGGAAGCTGAGCTTCAAGAGAGCGCTTCT 878

1441 AATTAACCGCTTCCCTTGGCCCAAAAGTCTGAAGATTTCTTATAAAGCTAGACTTTT 1500
 879 CATCAAGCTCGGCGCAAGATGCAATAGTCCGTACCAAGATCTCTGGAATCTTATGCG 938
 1501 CATCAAGATTCGGCTTGAGAGCAAGACAGATGAAAGTACCAATCGGATTTCAACTTCC 1560
 939 CAGTGGGATTTCTGCAAGTCTCTTGTGAAATCTGTGTGAAACATCATGCTCTTCTTGA 998
 1561 CAGTTACGAGACAGCTAAGAAATTAAGAAAGTGTGTGAAACATCAACGTTTTCAG 1620
 999 ACTTTTGAAGACCCAAACCAAGCCCAAGCCGCTCTCTTTAGCCGGGGTCAATCATT 1058
 1621 A---TTGACATCTACAGACACCATTTCCCAAAAGCAAAATTTCTGCGCTAGATCCAAAT 1677
 1059 TCGGTTGAGTGTGGACTCAGAAAGCAG 1086
 1678 TCGATACAGTGGCGGACTCAAGCTCAG 1705

RESULT 3
 US-09-664-958-5
 ; Sequence 5, Application US/09664958
 ; Patent No. 6916912
 ; GENERAL INFORMATION:
 ; APPLICANT: Trakht, Ilya
 ; APPLICANT: Canfield, Robert
 ; APPLICANT: Kalantarov, Gary
 ; APPLICANT: Rudchenko, Sergei
 ; TITLE OF INVENTION: No. 6916912el Tumor-Associated Marker
 ; FILE REFERENCE: 0575/60240
 ; CURRENT APPLICATION NUMBER: US/09/664,958
 ; CURRENT FILING DATE: 2000-09-18
 ; NUMBER OF SEQ ID NOS: 28
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 5
 ; LENGTH: 6263
 ; TYPE: DNA
 ; ORGANISM: Human
 ; US-09-664-958-5

Query Match 5.2%; Score 180.4; DB 3; Length 6263;
 Best Local Similarity 52.9%; Pred. No. 1.9e-37;
 Matches 485; Conservative 0; Mismatches 416; Indels 15; Gaps 4;

185 TGCTGATGACACCAAGGACATTTGAAGTTCCACAAAGAGCTCTGGAAGGTCTGC 244
 469 TGCTGATGCTCTGGAGATGATGATGAGTGTGAGAGAAATATGCGGGCCAGGTGCTGT 528
 245 TGATGACGTTTGCACCAACCTCAACCTGTGGAAGGTGACTATTTTGGCTCGAGTTTC 304
 529 TTGACCTGTCTGTGACCACTCAACCTCTAGAGAAAGACTTCTGCGCTGACCTTCT 588
 305 CTGATCACAAAAAGATCAAGGTGTGCTGTGATCTCTTAAACCCATTGTGAAACAGATTA 364
 589 GTGATGCTGACAGCCAGAAAGACTGTGAGACCCCTCCAAAGAGATTAAGAGAGATCC 648
 365 GAAGGCCAAAGACGTTGTGTTAAGTTTGTGTAAGTTCTTTCGCGCTGACACACAGC 424
 649 GAG---TAGCCCTGGAATTTTGTCTTACAGTCAAGTTTACCGCGCTATCTGTGCC 705
 425 AACTCCAAAGAACTCAAGAGTAACCTGTTCCGCTGACAGGTGAAGAGAGACTTGGCTC 484
 706 AGCTGACAGAAAGATCAAGATTAATCTGTGCTGCTGACGCTGCGGAGACATCATCA 765
 485 AAGGAGGTTGACGTGTATATACACACAGCGAGCTCTTGTATTTCAACATTTGTGCAAT 544
 766 CGGGCGGCTCCATGCTCTTGTGACGATGCGCTTACTGAGCTTCAACCTGTGCGC 825
 545 CTGAGATTGGGAAATTTGATGAAGCTTGAACAGAGACCTTGAAGCAAAAATAAATATACA 604
 826 CTGAGCTGGGTGACTATGATGCTGAGAGCATGTGGGCAACTATATGTCAAGCAGCTCGCT 885

QY 605 TACC-----TCAGCAAGACGCACTAGAGGACAAATCGTGAATTTACCATTAACCA 658
DB 886 TCCCCCTTAACCAAGACCCCGGAGCTGAGAGAGATCATGAGCTGCAATAGACATATA 945
QY 659 TTGGACAAACACGACGAGAAATCGAATTTCCAGCTCTTAGAATTTCCCGCTAGAGA 718
DB 946 GGGGGATGACCCCGGAGAGAGAAATCCACTTTAGAGAAATGCCAAGAGCTTTCCA 1005
QY 719 TGTATGAAATCCCGTTTGCACCCCGCCCAAGAGACAGGAGGACGAAATCTGCGCG 778
DB 1006 TGTACGAGATGACCTGACACATGCTCAAGAGACTTGAGGGCATCGAATCATGTTAGCG 1065
QY 779 TTGCAACAACGCGAAATCTAGTGTTCAGGGTTTCACTAGATCAATGCTTCAACTGAG 838
DB 1066 TTGTGTCATATGCTCTGCTCATCTACCGGAGCCGGCTGAGAAATCAACCGCTTGTGCTGCG 1125
QY 839 CCAAGGTGCGGAAGCTGAGCTTCAAGAGAAAGCGCTTTCTCATCAAGCTCCGCGCAAGT 898
DB 1126 CCAAGATCTCTCAAGATCTCTTACAGAGAGATTAATCTTATCAAGATCCGGCTGAGG 1185
QY 899 CCAATAGTGTGCTACAGAAATCTTGAATCTTGAATGCGCAATCTCGCAAT 958
DB 1186 AGTATAGCAATTTGAGAGACAAATGCTTTAAGCTCCCAACCAACCGGCTCAGCCAA 1245
QY 959 CCTTCTGAAATCTGTGTTGAACATCATGCTTCTTGAAGAGCCCAAC 1018
DB 1246 GACTGTGGAAGGTCTGATGAGATCATATTTCCGCGC---TGTGTCTCCCTGAGC 1302
QY 1019 CAAAGCCCAAGCCCGTCTCTTTAGCCGGGGTCAATCATTTCCGTTCAAGTGTGAGCTC 1078
DB 1303 CCCCACCCCAAGGCTTCTGTGTA---TGGGCTCAAGTTCCGGTACAGTGGAGAGGCC 1359
QY 1079 AGAAGCAGGTTCTCGA 1094
DB 1360 AGGCACAGACTCGCA 1375

RESULT 4

US-09-566-921-57
; Sequence 57, Application US/09566921
; Patent No. 668288
; GENERAL INFORMATION:
; APPLICANT: Loring, Jeanne F.
; APPLICANT: Tingley, Debra W.
; APPLICANT: Edwards, Carla M.
; TITLE OF INVENTION: GENES EXPRESSED IN ALZHEIMER'S DISEASE
; FILE REFERENCE: PA-0024 US
; CURRENT APPLICATION NUMBER: US/09/566,921
; CURRENT FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 138
; SOFTWARE: PERL Program
; SEQ ID NO 57
; LENGTH: 6268
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6682888 453592.3
; LOCATION: 2956, 4230-4252, 4825-4887
; OTHER INFORMATION: a, t, c, g, or other
US-09-566-921-57

Query Match 5.2%; Score 180.4; DB 3; Length 6268;
Best Local Similarity 52.9%; Pred. No. 1.9e-37;

Matches 485; Conservative 0; Mismatches 15; Gaps 4;

QY 185 TGTGTGATACACCCGAGGAGCATTTGAAGTTCCAAAGAGAGCTCTGGAAGAGCTCTGC 244
DB 469 TGTGTATGCTTGTGAGTATGAGTGTGAGGAAACATGCGCGGCGAGGTCTGT 528
QY 245 TGTATCAGTTTGCAACCACTCAACCTGCTGAGAGGTGATTTTGTGCTGAGTTTC 304

DB 529 TTGACCTGATCTGTGAACACTCAACTCTCTAGAGAGACTACTTCCGCTGACCTTCT 588
QY 305 CTATATCAAAAAGATCAAGGTGTGCTGATCTCTTAAACCATTGTGAACAGATT 364
DB 589 GTATGCTGACAGCCAGAAAGATCTGCTGACCCCTCCAAAGAGATCAAGAGACATCC 648
QY 365 GAAGGCCAAGACGTTGTTGTTAAGTTGTGTGAATTTCTTCCGCTGACACACAC 424
DB 649 GAG---TACCCTTGAATTTTCCCTTACACTCAAGTTTCTACCCGCTGATCTGCCC 705
QY 425 AACTCAAGAGAACTCAAGAGTAATCTTTCGCTGCAAGGTGAAGAGACTTGGCTC 484
DB 706 AGCTGACAGAAAGCATCAAGATTAATCACTGAGCTGAGAGTGGAGAGATCATCA 765
QY 485 AAGGCGAGTTGACGTGTAATGACACAGGAGCTCTCTTGAATTTACATTTGTGCAAT 544
DB 766 CCGGCGGCTGCTCATCTCTTGTGACAGATCCCTTACGAGGCTCTTCCGCTGACAG 825
QY 545 CTGAGATGAGGAGTTTGTGAAGACCTTGGACAGAGCACTTAGCAAAAATTAATTA 604
DB 826 CTGAGCTGAGTACTATGATCTGAGAGAGATGTGGCAATATGTACAGAGCTCCGCT 885
QY 605 TACC-----TCAGCAAGACGCACTAGAGACAAATCGTGAATTTCACTTAACACA 658
DB 886 TCCCCCTTAACCAAGACCCCGGAGCTGAGAGAGATCATGAGCTGCAATAGACATATA 945
QY 659 TTGGACAAACACGACGAGAAATCGAATTTCCAGCTCTTAGAATTTCCCGCTAGAGA 718
DB 946 GGGGGATGACCCCGGAGAGAGAAATCCACTTTAGAGATGCAAGAGCTTTCCA 1005
QY 719 TGTATGAAATCCCGTTTGCACCCCGCCCAAGAGACAGGAGGACAGAAATCAATCTGAGCG 778
DB 1006 TGTACGAGATGACCTGACACATGCTCAAGAGACTTGAGGGCATCGAATCATGTTAGCG 1065
QY 779 TTGCAACAACGCGAAATCTAGTGTTCAGGGTTTCACTAGATCAATGCTTCAACTGAG 838
DB 1066 TGTACGAGATGACCTGACACATGCTCAAGAGACTTGAGGGCATCATATGTTAGCG 1065
QY 719 TGTATGAAATCCCGTTTGCACCCCGCCCAAGAGACAGGAGGACAGAAATCAATCTGAGCG 778
DB 1006 TGTACGAGATGACCTGACACATGCTCAAGAGACTTGAGGGCATCATATGTTAGCG 1065
QY 779 TTGCAACAACGCGAAATCTAGTGTTCAGGGTTTCACTAGATCAATGCTTCAACTGAG 838
DB 1066 TTTGTGCAATGAGCTGCTCATCTACCGGAGCCGGCTGAGAAATCAACCGCTTGTGCTGCG 1125
QY 839 CCAAGGTGCGGAAGCTGAGCTTCAAGAGAGAGCGCTTTCTCATCAAGCTCCGCGCAAGT 898
DB 1126 CCAAGATCTCTCAAGATCTCTTCAAGAGAGATTAATCTTATCAAGATCCGGCTGAGG 1185
QY 899 CCAATAGTGTGCTACAGAAATCTTGAATCTTGAATGCGCAATCTCGCAAT 958
DB 1186 AGTATAGCAATTTGAGAGACAAATGCTTTAAGCTCCCAACCAACCGGCTCAGCCAA 1245
QY 959 CCTTCTGAAATCTGTGTTGAACATCATGCTTCTTGAAGAGCCCAAC 1018
DB 1246 GACTGTGGAAGGTCTGATGAGATCATATTTCCGCGC---TGTGTCTCCCTGAGC 1302
QY 1019 CAAAGCCCAAGCCCGTCTCTTTAGCCGGGGTCAATCATTTGGTTCAAGTGTGAGCTC 1078
DB 1303 CCCCACCCCAAGGCTTCTGTGTA---TGGGCTCAAGTTCCGGTACAGTGGAGAGGCC 1359
QY 1079 AGAAGCAGGTTCTCGA 1094
DB 1360 AGGCACAGACTCGCA 1375

RESULT 5

US-09-949-016-1488
; Sequence 1488, Application US/09949016
; Patent No. 6812338
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768

/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FASTSEQ for Windows Version 4.0
/ SEQ ID NO 1488
/ LENGTH: 4292
/ TYPE: DNA
/ ORGANISM: Human
US-09-949-016-1488

Query Match 4.9%; Score 169.6; DB 3; Length 4292;
Best Local Similarity 51.4%; Pred. No. 1.2e-34;
Matches 472; Conservative 0; Mismatches 434; Indels 12; Gaps 3;

```
175 AAAAAACAGATGTGGATGACACCCGAGGAGCATTTGAAGTTCCACAAGAGCTCTGGG 234
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 675 AAGATGACCCCTCTTAATGACACCGAATACAGCTGTGACCTGGAGAAACATGCCAAGGA 734

235 AAGTGTCTGTGATGACATTTGCAACCACTCAACCTGTGAGAGTACTATTTGGC 294
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 735 CAAGTGTATTGACAAAGTGTGTGAACCTCAATCTTTGAGAAAGACTACTTTGGA 794

295 CTCGAGTTTCTGTATCAAAAAGATCAGGTGTGCTGATCTCTAAACCATTGTG 354
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 795 CTTTGTGTTTCAGGAAAGCCCTGACGAAAACTGGTTAGATCTGCTAAAGAAATTAAG 854

355 AAACAGATTAGAAAGCCAAAGACGTTGTTAATTTGTGTGAATTTCTTCGGCT 414
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 855 AGACAACTAGAAAC--TTCCATGGCTATTCACTTTAATGTGAAGTTTATCTCTCT 911

415 GACCAACACAACTCCAAAGAAAGTCAACAAGTACTGTGGCTGACGATGAAGCAG 474
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 912 GATTCCTTCTCAATGACTGAAGATATACCAAGATATCTTGTGCTTCAAGCTCGGAG 971

475 GACTTGTCTCAAGGAGGTTGACGTGTAATGACACCAAGGACGCTCTTTGATTTCAAC 534
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 972 GACATTTGCTCTGGGCGCTGCTGCTCTGTTGTGATCATGATCTCTCGGAGATCTAC 1031

535 ATTGTGCATCTGAGATTGGGGATTT-----GATGAAGCCTTGGACAGAGACCTTA 588
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1032 ACCCTGACAGCTGAACTTGTGTGACTATGACCCAGAAAGAAATGACGACTCGACTCACT 1091

589 GCAAAAAATTAATATACCTACCTGACAAAGACGCACTAGAGAGCAAAATCGTGAATTTCA 648
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1092 GAATTCAGTTTGGCCCTTATCGCACTAAGAGCTGAAAGAGAGGTGACAGGCTGCAC 1151

649 CATTAACCATTTGACAAACACCAAGCAATCAGATTTCCAGCTCTTAAGATTGCCGT 708
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1152 AAAACCCACAGGGGGCTTATCGCCACAGCAAGCTGATTCACATGTTCTTAAGAAATGCAAG 1211

709 CGGCTAGAGATGATGGAATCCGGTTGCAACCGGCGCAAGGACGAGAGCAAGCAAGATC 768
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1212 AGGCTTTTCATGATGATGTGTTGACCTACATCATGCGCAAGACTCAGAAAGGTGAGCATC 1271

769 AATCTGCCGTTGCCAAACGAGGAATTTCTAGTGTTCAGGGTTTCACTAAGATCAATGCC 828
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1272 AAGCTGGGGGTGTGTCTAATGGAATTTCTCATTTTCAAAAGACAGACTGGGAATCATGCT 1331

829 TTCACTGGGCGCAAGGTGCGGAAGCTGACGTTCAAGAGAGAGCGCTTTCTCATCAAGCTC 888
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1332 TTTGCTTTGCGCAAAATCTTAAAAATTTCTTAAACGCAATCACTTACATTAAGATC 1391

889 CGGCGAGATGCCAATAGTGCTACAGGAATCTTGAATTCCTGATGGCGAGTGGCGAT 948
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1392 AAGCCGCGAGAGCTGGAACAGTTTGAAGATCCATTGATTCAAACTGCCAAACACCGG 1451

949 TTCTGCAAGTCTTCTTGAAAAATCTGTGTGAAACATCATGCTCTTTTGAATTTTGA 1008
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1452 GGAAGGAAAAAGACTAAGAAAGTGTGTGAGCATCATCTTTTCAAGGCTGTGTT--- 1508

1009 GAGCCCAACCAAGCCGCTCTTTAGCCGGGGGTCACTATTTGGGTTCACT 1068
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
Db 1509 TCTCCAGAGCACCCCAAAAGCAAGTTCTGACTTGGGGTCCAAATTTGCTAATG 1568
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1069 GGTGGACTCAGAAAGCAG 1086
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1569 GGCCGACCCACAGACAG 1586
```

RESULT 6
US-09-949-016-208
Sequence 208, Application US/09949016
Patent No. 6812339

GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CLO01307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 208
LENGTH: 4336
TYPE: DNA
ORGANISM: Human
US-09-949-016-208

Query Match 4.9%; Score 169.6; DB 3; Length 4336;
Best Local Similarity 51.4%; Pred. No. 1.2e-34;
Matches 472; Conservative 0; Mismatches 434; Indels 12; Gaps 3;

```
175 AAAAAACAGATGTGGATGACACCCGAGGAGCATTTGAAGTTCCACAAGAGCTCTGGG 234
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 705 AAGTGAACCCCTCTTAATGACACCGAATACAGCTGTGACCTGAGAAACATGCCAAGGA 764

235 AAGTGTCTGTGATGACATTTGCAACCACTCAACCTGTGAGAGTACTATTTGGC 294
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 765 CAAGTGTATTGACAAAGTGTGTGAACCTCAATCTTTGGAAGAAAGACTACTTTGGA 824

295 CTCGAGTTTCTGATCAAAAAGATCAGGTGTGCTGATCTCTTAAACCATTGTG 354
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 825 CTTTGTGTTTCAGAAAGCCCTGAGCAAGAAATCTGTAAATCTGCTAAAGAAATTAAG 884

355 AAACAGATTAGAAAGCCAAAGCACTGTGTTAAAGTTGTGTGAATTTCTTCGGCT 414
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 885 AGACAACTGAGAAAC--TTCCATGGCTATTCACTTTAATGTGAAGTTTATCTCTCT 941

415 GACCAACACAACTCCAAAGAAAGTCAACAAGTACTGTGGCTGACGATGAAGCAG 474
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 942 GATTCCTTCTCAATGACTGAAGATATACCAAGATATCTTGTGCTTCAAGCTCGGAGC 1001

475 GACTTGTCTCAAGGAGGTTGACGTGATGACACCAAGGACGCTCTTGAATTTCAAC 534
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1002 GACATTTGCTCTGGGCGCTGCTGCTCTTTTGTGATCATGCTCTCTGGGATCTTAC 1061

535 ATTGTGCATCTGAGATTGGGGATTT-----GATGAAGCCTTGAACAGAGACCTTA 588
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1062 ACCCTGACAGCTGAACTTGTGTACTATGACCCAGAAAGAAATGAGCATGCACTCACT 1121

589 GCAAAAAATTAATATACCTCAGCAAGACGCACTAGAGGACCAAAATGCTGAATTTAC 648
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1122 GAATTCAGTTTGGCCCTTACTAGACTAAGAGAGCTGAAAGAGGTGACAGCTGCAC 1181

649 CATTAACCATTTGACAAACACCAAGCAATCAGATTTTCAAGCTCTTGAATTTGCCCT 708
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1182 AAAACCCACAGGGGGCTTATTCGCAACAGCAAGCTGATTTCCAGTTCTTGAAGAAATGAAG 1241

709 CGGCTAGAGATGATGGAATCCGGTTGCAACCGGCGCAAGGACGAGAGCAAGCAAGATC 768
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
Db 1242 AGGCTTCCATGATGATGTTGACCTACATCATGSCCAAGGCTCCAGAAAGGTGGACATC 1301
Qy 769 AATCTGGCCCTTGGCAACACAGGGAATCTAGTGTTCAGGGTTTCACTAAGATCAATGCC 828
Db 1302 AACCTGGGCGTGTGTGTCTAATGACCTTCTATTAACAAGACAGACTGCCAATCAATCGT 1361
Qy 829 TTCACTGGGSCCAAGGTGGCGAAGCTGAGCTTCAAGAGGAAGGCGTTTCTCATCAAGCTC 888
Db 1362 TTTCCTTGGCCGCAAAATCTTAAATTTCTTAAACGCACTAATCTTCTCATTTAAAGTC 1421
Qy 889 CGGCCAGATGCCAATAGTGGTACAGAGATACCTTGAATTCCTGATGCGCACTCGGAT 948
Db 1422 AGACCGGCGAGCTGGAAAGTTGAGATACCATGATTTCAACGCAACACCGG 1481
Qy 949 TTCTGCAAGCTCTTGGAAAAATCTGTGTGGAACATATGCTTCTTTAAGCTTTTGA 1008
Db 1482 GCAGCGAAGAAAGATATGAAAGTGTGGTGAAGCATCATCTTCTAAGGCTTGT--- 1538
Qy 1009 GAGCCCAACCAAGCCCAAGCCGCTCTTTAGCCGGGGTTCATCTTTCGGTTCAGT 1068
Db 1539 TCTCCAGAGAGCACCACCAAGCTTCTGACCTTGGGGTTCCAATTTCCCTATAGT 1598
Qy 1069 GGTGGAATCAGAACGAG 1086
Db 1599 GCGCCGACCCCAAGCAG 1616
```

```
RESULT 7
US-09-270-767-1995
; Sequence 1995, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1995
; LENGTH: 576
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-1995
```

```
Query Match 4.8%; Score 165.8; DB 3; Length 576;
Best Local Similarity 61.3%; Pred. No. 4.7e-34;
Matches 285; Conservative 0; Mismatches 177; Indels 3; Gaps 1;
```

```
Qy 646 CACCATTAACCACTTGGACAAACACGACGATTCAGATTTCCAGTCTCTAGATTTGCC 705
Db 6 CACTAAAGATGTGTGTCATCCCGCGGAGCGGACCTTAACCTTGGAGAGCGCT 65
Qy 706 CGTGGCTAGAGATGTATGGAATCCGGTTCACCCGCGCAAGACAGGAAAGGACGAAAG 765
Db 66 CGAGGTGTGAGCTGTATGCGATGAAATGATCCGCGCAAGATGTGGAAAGGGGTCCG 125
Qy 766 ATCAATCTGGCCGTGGCAACACGGAATTTCTAGTGTTCAGGGTTTCACTAAGATCAAT 825
Db 126 CTTAATCTGGCTGTGGCCCATGCGGATCACAGTCTTTCAGAAACATCAAGCGATCAAC 185
Qy 826 GCCTTCACTGGGCAAGGTGGGAAAGCTGAGCTTCAAGAGAAAGGCGTTTCTCATCAAG 885
Db 186 ACCTTCTCGGGCTAAGATACGAGATTTCTTCAAGGCAAGGATTTCTGTCTAAA 245
Qy 886 CTCGCGCAGATGCCAATAGTGGTACAGAGATACCTTGAATTCCTGATGCGCACTCG 945
Db 246 CTCGATCCGAGGAGATATGATATTACAGAGATACGATGAGTTCTTTCGAGGGTCCG 305
Qy 946 GATTTCTGCAAGTCTTCTGGAAAAATCTGTGTGGAACATATGCTTCTTTAAGCTTTT 1005
Db 306 AACGAGTGCAAAAATTTCTGGAAAAATGCGTGCAGAAATCAAGATTTCTTCGATGCAC 365
```

```
Qy 1006 GAAGAGCCCA---ACCAAGCCCAAGCCCGTCTCTTAAAGCCGGGGTTCATCTTCCG 1062
Db 366 GCGCTACAAATATACGCCCGCAAGCCCAAACTCGGCTTCTTCGCGGGATGATTCGCCG 425
Qy 1063 TTCAAGTGTGCGACTGAGAAAGAGCTTCTGACTATGTATTAAGAA 1107
Db 426 TATAGCGAAAAACCCAGAAAGAGATTTATGAGTTGGTTCGGAA 470
```

```
RESULT 8
US-09-270-767-17277
; Sequence 17277, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 17277
; LENGTH: 576
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-17277
```

```
Query Match 4.8%; Score 165.8; DB 3; Length 576;
Best Local Similarity 61.3%; Pred. No. 4.7e-34;
Matches 285; Conservative 0; Mismatches 177; Indels 3; Gaps 1;
```

```
Qy 646 CACCATTAACCACTTGGACAAACACGACGATTCAGATTTCCAGTCTCTAGATTTGCC 705
Db 6 CACTAAAGATGTGTGTCATCCCGCGGAGCGGACCTTAACCTTGGAGAGCGCT 65
Qy 706 CGTGGCTAGAGATGTATGGAATCCGGTTCACCCGCGCAAGACAGGAAAGGACGAAAG 765
Db 66 CGAGGTGTGAGCTGTATGCGATGAAATGATCCGCGCAAGATGTGGAAAGGGGTCCG 125
Qy 766 ATCAATCTGGCCGTGGCAACACGGAATTTCTAGTGTTCAGGGTTTCACTAAGATCAAT 825
Db 126 CTTAATCTGGCTGTGGCCCATGCGGATCACAGTCTTTCAGAAACATCAAGCGATCAAC 185
Qy 826 GCCTTCACTGGGCAAGGTGGGAAAGCTGAGCTTCAAGAGAAAGGCGTTTCTCATCAAG 885
Db 186 ACCTTCTCGGTGCTAAGATACGAGATTTCTTCAAGCGCAAGGATTTCTGTCTAAA 245
Qy 886 CTCGCGCAGATGCCAATAGTGGTACAGAGATACCTTGAATTCCTGATGCGCACTCG 945
Db 246 CTCGATCCGAGGAGATATGATATTACAGAGATACGATGAGTTCTTTCGAGGGTCCG 305
Qy 946 GATTTCTGCAAGTCTTCTGGAAAAATCTGTGTGGAACATATGCTTCTTTAAGCTTTT 1005
Db 306 AACGAGTGCAAAAATTTCTGGAAAAATGCGTGCAGAAATCAAGATTTCTTCGATGCAC 365
Qy 1006 GAAGAGCCCA---ACCAAGCCCAAGCCCGTCTCTTAAAGCCGGGGTTCATCTTCCG 1062
Db 366 GCGTACAAATAAGCGCCGAGGCAAACTCGGTTCTCTCGGGGGTATGATTCGCCG 425
Qy 1063 TTCAAGTGTGCGACTGAGAAAGAGTTCTGACTATGTATTAAGAA 1107
Db 426 TATAGCGAAAAACCCAGAAAGAGATTTATGAGTTGGTTCGGAA 470
```

```
RESULT 9
US-09-848-294-1
; Sequence 1, Application US/09848294
; Patent No. 6479640
; GENERAL INFORMATION:
; APPLICANT: Tonks, Nicholas K.
; TITLE OF INVENTION: Isolation of A cDNA Encoding A No. 6479640e1
; TITLE OF INVENTION: Protein Tyrosine Phosphatase Which Localizes to Focal
```

```

; TITLE OF INVENTION: Adhesions and Uses Therefor
; FILE REFERENCE: CSHL90-04FZA
; CURRENT APPLICATION NUMBER: US/09/848,294
; CURRENT FILING DATE: 2001-05-03
; PRIOR APPLICATION NUMBER: 09/235,251
; PRIOR FILING DATE: 1999-01-22
; PRIOR APPLICATION NUMBER: 08/759,536
; PRIOR FILING DATE: 1996-12-04
; PRIOR APPLICATION NUMBER: 08/107,420
; PRIOR FILING DATE: 1993-08-16
; PRIOR APPLICATION NUMBER: 07/663,579
; PRIOR FILING DATE: 1991-03-01
; PRIOR APPLICATION NUMBER: 07/494,036
; PRIOR FILING DATE: 1990-03-14
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 3984
; TYPE: DNA
; ORGANISM: Homosapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (24)...(2765)
; US-09-848-294-1

```

```

Query Match      4.2%; Score 144; DB 3; Length 3984;
Best Local Similarity 50.8%; Pred. No. 8.5e-28;
Matches 423; Conservative 0; Mismatches 400; Indels 9; Gaps 3;

```

```

OY 174 CAAAATCCAGATGCTGATGACACCCAGAGGCGATTGAGTTCCACAAAGAGCTCTGG 233
DB 116 GAGCATCCACTTTTATGATGGCGTGTACAGACCTTAAAGTTACTTAACAGACACTGG 175
OY 234 GAAGTGTCTGCTGATGATGCACTTGGCAACCACTCAACCTCGGGAAGGAGTCTATTGG 293
DB 176 CCAGGTTCTTGTGATGATGTCACAAACCACTGGGTGACTGAAGAAATATTTGG 235
OY 294 CCTCGAGTTTCTGATCACAAAAAGATCAAGTG--TGCTGATCTCTTAAACCAT 350
DB 236 TTTCAGCATGATGATGACGATCCGCTGATCTCTTATGATGCTGGAAGCAAGCAACCAT 295
OY 351 TGTGAAACAGATTAGAAAGGCCAAAGCAGTGTGTTGTAAGTTGTGTAATTTCTTCC 410
DB 296 CAGGAAGCAGTTAAAGAGAGTTTCCCTGTACCTGACATTTTCAGTAAAGATTTTAT 355
OY 411 GCTGACCAACAACAACCTCCAGAGAACTACACAGTATCTGTCGCTGAGGTGA 470
DB 356 ACCTGATCCCAACAACCTGACAGCAAGAAACAAACAGGCACTGTATTTCTTCAACTGA 415
OY 471 GCAGGACTGGCTCAAGGAGGTTGACGTGTAATGACACAGGAGCACTCTTGAATTC 530
DB 416 GATGATATTTGGAGAGAAAGTTAACCTGCCCTTAACTCAGCAGTGGTCTTAGCGTC 475
OY 531 ACACATTTGCAATCTGAGATTGGGATTTTGAATGAAGCTT--GACAGAGAGCACTT 587
DB 476 CTATCCCGTACATCTCATTTTGGAGACTATATTTCTTCATACATCATCCAGGCTATCT 535
OY 588 ACCAAAAAATAATACATACCTCAGCAAGACGACACTAGAGCAAAATGTGGAATTTCA 647
DB 536 TTCCGATGATCACTTTATACCCGATCAAAATAGAGACTTTTAAACAAAATGGAATCTT 595
OY 648 CCAT---AACCAATTTGACAAACACACAGAGATCAGATTTCCAGCTCTAGAGATTGC 704
DB 596 GATGAGCAGCAGCAGTGGGCTTAAACATCAGAGAGAAATCTCTGTATATCAACTAGC 655
OY 705 CGGTGGCTAGAGATGATGAAATCGGTTGACCCGGCCAAAGACAGGAGAGGACGAA 764
DB 656 GGGAGCCCTCGACTTCTATGAGTAGAAGCTGCAAGATGAGGATCTGCAATTTAGA 715
OY 765 GATCAATCTGGCGCTGGCCAAACGGGAATTTCAAGTTTTCAGGGTTTCACTAAGTCA 824
DB 716 CCTAATGATTTGAATTTGCTTCGGGGGTGTGTGTGACGAAAATATCATTTGCAAG 775

```

```

OY 825 TGCCTTCACTGGGCGCAAGTGCAGAGCTGAGCTTCAAGAGAAAGCCCTTTCATCA 884
DB 776 TTTCATCTCTGGGAGAACATTTCTCAAAATTTCTTCAAAAGAAAGTTCTTATACA 835
OY 885 GCTCCGCGCAAGTGCCTAATGATGCTGACAGAGATCTTGAATTTCTGATGCGCACTG 944
DB 836 TCAGCGACAGAAACAGGCTGAATCCAGGAACTATTTGCGCTTCAACATGCTGAATTA 895
OY 945 GGATTTTCAAGTCTTCTGAAATCTGTGTTGAATCATCATGCTTTT 996
DB 896 CCGATCTTGCAAAAATTTGGAATCTGTGTGAGCACCATATGTTCTTT 947

```

```

RESULT 10
US-09-906-779-3
; Sequence 3, Application US/09906779
; Patent No. 6770466
; GENERAL INFORMATION:
; APPLICANT: Shi et al.
; TITLE OF INVENTION: Human Protein Tyrosine Phosphatase Polynucleotides, Polypeptides,
; FILE REFERENCE: PTO4091
; CURRENT APPLICATION NUMBER: US/09/906,779
; CURRENT FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: PCT/US01/01563
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/176,306
; PRIOR FILING DATE: 2000-01-18
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2872
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-906-779-3

```

```

Query Match      4.2%; Score 143.2; DB 3; Length 2872;
Best Local Similarity 50.3%; Pred. No. 1.2e-27;
Matches 490; Conservative 0; Mismatches 468; Indels 16; Gaps 5;

```

```

OY 131 GACGAGACCGGCCCAACACCTTCAGGAAAATCTGTGCTCATCAAAATCCAGATGCTGG 190
DB 1 GGCAGAGACCGCGCGCGCGCGCCAGAGCCAC-CCCTTACTGCGCGCTTCTGCTGCG 59
OY 191 ATGACACCGAGAGGCAATTGTAAGTTCCACAAAGAGCTCTGGAAGGCTGCTGATG 250
DB 60 ACGGACCGAAGTAGCGGTGACCTGCGCAAAACATGCCAAAGGCAAGATTTGTTGATC 119
OY 251 CAGTTTGCAACCACTCAACCTCGTGAAGGTGACTATTTTGGCCTGAGTTTCTGATC 310
DB 120 AGATTGTATCACTTGAACCTTGTGAAACGATTAATTTGGCTTCCAGTTCTGCACT 179
OY 311 ACAAAAAATACAGGTGTGCTGATCTCTTAAACCATTTGTGAACAGATTAAAGGC 370
DB 180 CTGCCAGGTGCGGACGTGATGATGCAACCAACCATATAAAAGAGATGAATAATGG 239
OY 371 CAAGCAGTGTGTTGTAATTTGTGTAATTTCTTCCGCTGACACACACAACACTCC 430
DB 240 GACCTGCTTAATCTTAACTTTGAGTTAAATTAATTAATTTCTTCAAGAACCAACACTTC 299
OY 431 AAGAGAACTCAACAGATACCTGTTGCGGCTGACAGTGAAGAGGACTTGGCTCAAGCA 490
DB 300 GTGAGAGATTAAACAGATACCTGTTGTTTAACTCAGGCAATGATCAATCTTCTGGA 359
OY 491 GGTACGTGTAATGACACACAGCGAGCTCTTGAATTTCAACATTTGTGCAATCTGGA 550
DB 350 AATTGAAATGCGCTTATGAAACAGCTGTGAATTAAGCTCTCTGTCTACAAAGGAGAC 419
OY 551 TTGGGATTTTGAATGAAGCTTTGACA---GAGAGCACTTGAACAAAATAATATATATAC 607
DB 420 TTGGGAGGTGAGGCTTTCAGAAACACACACAGAGCTTGTGTAGATTTTGGGTTCTTC 479
OY 608 C---TCAGCAAGACGCACTAGAGCAAAATCTGTGAATTTTCACATTAACCATTTGAGC 664

```

Db 480 CAATCAGACAGAGCAATGGAATTGATCTTCCAGAGATGAGAAAGTCAGGGGAA 539
 Qy 665 AAACACCAGCAGATCAGATTTCCAGCTCTTAAGATTTCCCGCTGAGATGATGATG 724
 Db 540 AGAGCCCTGCGCGAGGAGCTCTCTATCTGAATAAAGCGAATGAGTGAATATG 599
 Qy 725 GAATCCGGTTGACCGCGCCAGAGAGAGGAGCAGAAATCAATCTGCGCGTGGCA 784
 Db 600 GGGTAAACATGACGTTGTGAGGAGAGAGATGCTGTGAATTTCTCTTGACCTGACC 659
 Qy 785 ACAACGGAATTTAGTGTTCAGGTTTCACTAAGATCAATGCTTCAACTGCGGCAAG 844
 Db 660 CAGCAGCATATTAATCTTTGAAGAGCTAAGAAATAGGCTTATCTTTGGCTTAAA 719
 Qy 845 TCGGGAAGCTGAGCTTCAAGAGAGAGCGCTTTCTCATCAAGCTCGGCGCAGATCCA 904
 Db 720 TTACCAAAATGATTTTAAAGAGCAATTAACACTGCTGTGTGATGATGATC 779
 Qy 905 -----GTGCGTACAGATACCTTGGAAATTCGTATGCGAGTTCGCGAATTCGAA 958
 Db 780 AGGAGCTAAGCAGAGACACGTTGTGTTCCGTTAGACAGTGCAGAGACTGCAAC 839
 Qy 959 CCTTGTGAAAATCTGTGTGAACATCATGCTTCTTGAAGTCTTTGAAGCCCAAC 1018
 Db 840 ACTTTGGAAGTGTGAGTTGAGCAGCAGCATCTTCCAGCTGCGGAGCGCCAGAAAC 899
 Qy 1019 CAAAGCCCAA---GCCGCTCTCTTTAGCCGGGGCTCATCATTTGCGTTAGTGTG 1075
 Db 900 GCAAAATCCAAATGATCCGACTTATCATGCGGTGCTCGCTTCAGATTCAGTGGCG 959
 Qy 1076 CTCAGAGCAGGTT 1089
 Db 960 CAGATATCAAGCT 973

RESULT 11
 US-09-949-016-3593
 ; Sequence 3593; Application US/09949016
 ; Patent No. 6812339
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 ; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CL001307
 ; CURRENT APPLICATION NUMBER: US/09/949,016
 ; PRIOR FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: 60/241,755
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/237,768
 ; PRIOR FILING DATE: 2000-10-03
 ; PRIOR APPLICATION NUMBER: 60/231,498
 ; PRIOR FILING DATE: 2000-09-08
 ; NUMBER OF SEQ ID NOS: 207012
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 3593
 ; LENGTH: 2156
 ; TYPE: DNA
 ; ORGANISM: Human
 US-09-949-016-3593

Query Match 4.1%; Score 142.4; DB 3; Length 2156;
 Beest Local Similarity 50.7%; Pred. No. 1.7e-27;
 Matches 422; Conservative 0; Mismatches 401; Indels 9; Gaps 3;
 Qy 174 CAAATTCAGATCTGATGATGACCAAGGAGGATTTGAAGTTCCAAAGAGCTCTGG 233
 Db 110 CAGCATCCATTTTATGATGCGCTGTGATGACACCTTTAAAGTTAAACAAAGACCTGG 169
 Qy 234 GAAGTGTCTGATGATGACGTTTGAACCACTCAACTGTGGAAGGTGACTATTTTGG 293
 Db 170 CCAGGTTCTTCTGATGATGATGATGACCAACCACTGGGTGTGATGAAAAGAAATATTTGG 229

Qy 294 CCTGAGTTTCTTGATCAGAAAAGATCAAGG---TGGCTGATCTCTTAACCCAT 350
 Db 230 TTACAGCATGATGACGATCCGCTGACCTTCTTAAGTGGCTGAGCAAGACCAT 289
 Qy 351 TGTAAACAGATTGAAGGCCAAGACGTTGTGTTAAGTTGTGATGAAATCTTTCC 410
 Db 290 CAGGAAGAGTTTAAAGAGGTTTCCCTGTACCTTCGATTTTTCGATGAAGATTTTAT 349
 Qy 411 GCTTACACACCAACTCCAGAGAACTCAGAGTACTTTTGGCGCTGAGGTGA 470
 Db 350 ACCTGATCCCAACACTCAGACAGCAACCAAGCAGCTTGTATTTCTTCAACATGAA 409
 Qy 471 GCAGACTTGGCTCAAGCAGAGTTGACGTATACACACAGCGAGCTCTGATTTTC 530
 Db 410 GATGATATTTGGAAGAGAGTTAACCTGCTCTTAATCTCAGCAGTGTCTTACGTC 469
 Qy 531 ACATTTGTGCAATCTGAAATTTGGGATTTTGAAGCTT---GAGACAGAGCACTT 587
 Db 470 CTATGCCGTACATCTCATTTTGAAGACTATATTTCTCATATCATTCAGGCTATCT 529
 Qy 588 AGCAAAAATTAATACATACCTCAGACAGCACTAGAGACAAATGTGAAATTTCA 647
 Db 530 TTCCGATAGTCACTTTATACCGATCAAAATGAGACTTTTAAACAAAGTCGAACTCT 589
 Qy 648 CCAT---AACCATTTGACAAACACACAGCAGATGCAATTCAGCTCTAGATTCG 704
 Db 590 GCATGACGACAGAGTGGCTTAACATCAGAGCAAGATCTGCTATATCAACTAGC 649
 Qy 705 CCGTCGCTAAGATGATGATCCGTTTGACACCGGCGCAAGAGAGGAGGACAGAA 764
 Db 650 GGGAGCCCTCGACTTGTATGAGATGAACTGACAGTGTGATGAGGATCTCACAATTA 709
 Qy 765 GATCAATCTGCGCTTCCCAACAGCGGAATTTAGTGTTCAGGTTTCACTAAGTCA 824
 Db 710 CCTAATGATGGAATTTGCTTCGCGGGTGTGTGTGTCGAAATAATACATTTGCAAG 769
 Qy 825 TGCCTTCACTGGGCGCAAGTGTGAGCTGAGCTTCAAGAGAGCGCTTCTCATCA 884
 Db 770 TTTCTATCTTGGAGGACATTTCTCAAAATTTCTTCAAGAGAAAGTCTTCTTACA 829
 Qy 885 GCTCGGCCAGATGCAATAGTGCCTACAGAGATACCTTGAATTCCTGATGCGCAGTGC 944
 Db 830 TCAGGACAGAAACAGGCTGAAATTCAGGGAATATTTGTGGCTTCAAGATGTGATTA 889
 Qy 945 GAATTTTGAAGTCTTCTGAAAATCTGTGTGAATCAATGCTTCTTT 966
 Db 890 CCGATCTTGCAAAAACCTGTGAAAATCTGTGTGAGCACCATACCTTCTTT 941

RESULT 12
 US-09-799-451-201
 ; Sequence 201; Application US/09799451
 ; Patent No. 6783969
 ; GENERAL INFORMATION:
 ; APPLICANT: Tang, Y. Tom
 ; APPLICANT: Zhou, Ping
 ; APPLICANT: Goodrich, Ryle
 ; APPLICANT: Asundi, Vinod
 ; APPLICANT: Ren, Feiyan
 ; APPLICANT: Zhang, Jie
 ; APPLICANT: Xue, Aidong J.
 ; APPLICANT: Zhao, Qing A.
 ; APPLICANT: Wang, Jian-Rui
 ; APPLICANT: Ma, Yungting
 ; APPLICANT: Yamazaki, Victoria
 ; APPLICANT: Chen, Rui-hong
 ; APPLICANT: Wang, Zhiwei
 ; APPLICANT: Wang, Dunrui
 ; APPLICANT: Yang, Yonghong
 ; APPLICANT: Wehrman, Tom
 ; APPLICANT: Ghosh, Reena
 ; APPLICANT: Drmanac, Radoje T.
 ; TITLE OF INVENTION: No. 6783969e1 Nucleic Acids and

FILE REFERENCE: 803
CURRENT APPLICATION NUMBER: US/09/799,451
CURRENT FILING DATE: 2001-03-05
NUMBER OF SEQ ID NOS: 948
SOFTWARE: pt_fl_genes Version 2.0
SEQ ID NO 201
LENGTH: 3398
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (174)..(2234)
US-09-799-451-201

Query Match 3.9%; Score 133.6; DB 3; Length 3398;

Best Local Similarity 49.5%; Pred. No. 4.8e-25;
Matches 495; Conservative 0; Mismatches 484; Indels 21; Gaps 5;

185 TGCTGGATGACACCCAGAGGCAATTGAAATTCACAAAGAGCTCTGGGAAGTGTCTGC 244
1155 |
319 TTTCTGATGTACTGATGTTAAGTGTGACCTTGCCAAAAAGCAAGAGAGTTGT 378
1069 |
245 TGGATGCACTTGGCAACCACTCAACCTCGTGAAGGTGACTATTTGGCTCGAGTTTC 304
1013 |
379 TTGATCAGATTATGTACCACTGGAACCTGATGAAAGCACTATTTGGCTGAGATTTA 438
1159 |
305 CTGATACAAAAAGATCAAGGTGTGCTGATCTCTTAAACCCATTGTGAAACAGATT 364
1199 |
439 TGGATTACAGACACAGTAGACATTTGGTGTGATGTAACAAAAAGCAAAAGCAAGTAA 498
1219 |
365 GAAGGCCAAGACAGTGTGTTGTTAAGTTGGTGAATTTCTTCCGCTGACCAACAC 424
1219 |
499 AATTGTGTTACCTTATTTGTCTGACATCTTGAGTTAAGTTTATCTCGAAGACCAATA 558
1219 |
425 AACTCCAGAAAGAACTCAGAAAGTACCTGTGCGCTGCAAGGTGAAGCAAGCTTGGCTC 484
1219 |
559 ACCTTGATGAGAGCTAACCCGGTATTTATTTGTTCTGAGTTAAACAAAGATATTCTCA 618
1219 |
485 AAGGCGAGTTGACGCTTAATGACACAGGAGAGCTCTTGAATTTACACATTGGCAAT 544
1219 |
619 GTGGAAATTTAAGCTGTCCCTTGTATACAGAGTCAATTTGGCAGTTAAATCTGCAAG 678
1219 |
545 CTGAGATTTGGGATTTTGAAGCCTTGGACAG--AGAGCACTTAGCAAAAAATAAT 601
1219 |
679 CTGAATTGTGATGATGATCTTGTGAGCATAAGTCTGAACTTGTCTAGAGTTCAAGT 738
1219 |
602 ACATACCC--TGACAAAGCGCACTAGAGGACAAATCGTGAATTTCAACATTAACACA 658
1219 |
739 TCGTGTCTATTCAGATGAAAGATGGAATCGGCTATTTTGAAGAAATGGAAGATATCA 798
1219 |
655 TTGACAAACACAGAGAGATCAGATTTCAGCTCTTAAGATGTCGCGCTGAGTAA 718
1219 |
799 GAGGTCAACACAGAGAGCTGAAACCAATTAATTAAGTAAAGCAATGAGCTAGAAA 858
1219 |
719 TGTATGGAATCCGGTTGACACCGGCCAAGACAGGAAAGCAAGATCAATCTGGCGG 778
1219 |
859 TGTATGGGTTGATATGATGATGCTGCAAGGCTAGAGATGGAATGATATGTTGGAGC 918
1219 |
779 TTGCAACACGAGGAAATTTAGTGTTCAGGGTTTCACTAAGATCAATGCTTCACTGGG 838
1219 |
919 TAAACACCAAGAGAGCTGTTTGTGTTTGAAGAGATACCAAAATTTGGCTTATTTTGGC 978
1219 |
839 CCAAGTGTGGGAAAGCTGAGCTTCAAGAGAGAGCGCTTTCATCAAGCTCCGCGCAGAG 898
1219 |
979 CCAAGTAAACAGAGTTGATTTTAAAGAAATAATTAACCTTGTGTGTGTAAGAAAGT 1038
1219 |
899 CCAAT-----AGTGGTACAGAGATACCTTGAATTTCTGATGCGCAGTCCGGAATTTCT 952
1219 |
1039 ATGATCAGGCGCAAGAAAGAGAAACATATTTGTCTTGAAGTGAATCAATCAAGAAAGAT 1098
1219 |
953 GCAAGTCTTTTGGAAATCTGTGTGAACATCATGCTTTCTTTAGACTTTTGAAGAGC 1012
1219 |

1099 GCAAACATTTATGAAATGTGCTGTGAGACATCATGCTTTTCCGCTTCAGAGCCCCG 1158
1013 CCAACCAAAAGCCCAAGC---CCGCTCTTTAGCCGGGGGATCATTTGGTTCAAGT 1069
1159 TCCAAAGAGTTTCTCATGATGATGATGATTTATTCGATGAGATCAGATTTAATATGATG 1218
1070 GTGGACTCAGAAAGAGGTTTCTGACTATGTTTAAAGAGAGACATTAAGAGTGCAGT 1129
1219 GGAACACAGATATGACACCAAAACCAATTAAGCAAGAAATCA-----CATCT 1272
1130 TTGAAAGAGACACAGCAAGATTCTTCTATCCGAGCT 1169
1273 TTGAAAGAGAGCCCGCAAGACGATATCTTACAGCAATCT 1312

RESULT 13

US-09-270-767-1542
Sequence 1542, Application US/09270767
Patent No. 6703491

GENERAL INFORMATION:

APPLICANT: Homburger et al.
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster

FILE REFERENCE: File Reference: 7326-094
CURRENT APPLICATION NUMBER: US/09/270,767

CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 62517

SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1542

LENGTH: 449
TYPE: DNA

ORGANISM: Drosophila melanogaster
US-09-270-767-1542

Query Match 3.1%; Score 106.4; DB 3; Length 449;

Best Local Similarity 65.7%; Pred. No. 3.4e-18;
Matches 155; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

2761 CCGCACCCGCGACACATGATGTCAGCTGTGTCGACCCGCAACAGCAGCTTCATG 2820
199 CCATCCCAAGAAACACCGCTGTGACGTCGTGTGCAATCGTGGCGCACCGTAGAGACTG 258
2821 GTGACCTCAGATGACAGTGAAGATCAATTTGTGAAACCTGCTGAGAAATTCAAA 2880
259 GCGCATCACTGATATCCGCGCAGACACAGATGTCGCGTATCTCTGGAATTAAG 318
2881 AACAGCAAGGCTGACAGAGCTGTGGGAGTGTCAAACTTGTGCTGCTTCTTAC 2940
319 AATACCTCGGCTGCAAAAGCTGTGGGAGTGTGACGTCCTTTGTGTGATCTTAC 378
2941 AATACACACAGCAATATCTCCCTTCCAGCTGCTGCTGCTGCTGCTGCTGCTGCT 2996
379 AAAACCTACAGAGAGAGTTCGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 434

RESULT 14

US-09-270-767-16824
Sequence 16824, Application US/09270767
Patent No. 6703491

GENERAL INFORMATION:

APPLICANT: Homburger et al.
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster

FILE REFERENCE: File Reference: 7326-094
CURRENT APPLICATION NUMBER: US/09/270,767

CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 62517

SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 16824

LENGTH: 449
TYPE: DNA

ORGANISM: Drosophila melanogaster
US-09-270-767-16824

Query Match 3.1%; Score 106.4; DB 3; Length 449;

	Best Local Similarity	65.7%	Pred. No. 3,4e-18	Matches 155	Conservative 0	Mismatches 81	Indels 0	Gaps 0
Qy	2761	CCGCACCGCGGCAACACATG	TGTCAGTGTGCGCACCGCAACACCAGGTTCCATG	2820				
Db	199	CCATCCCGACCAATACCGCT	CTGCACCTCTGTTGGCAATCGTGCGCCACCGTGAAGCTG	258				
Qy	2821	GTGACCTTCAGCATCGCAG	TGAGAGATCTGTCTGGAAACCTGCTAGGAATTCAAA	2880				
Db	259	GGCATTACCTGATATCCG	CCGAGACACAGATGTCGGTATCTGCTGCGAAATTCAG	318				
Qy	2881	AACGCAACGGGTGGCAGAA	CTGTGGGTGGTTCAAACTTCGTCTGTTCTTCTAC	2940				
Db	319	AATAGCTCCGGCTGGCAAA	AGCTCTGTGGGTGTTCAAGTCTTTTGTCTACTTTTAC	378				
Qy	2941	AATTCACACCGAGACATCA	TCCCTTGTCCACGCTGCTCTGTCTCGGCTACTGGCT	2996				
Db	379	AAAAGCTATCCAGACGAG	TTCGCACTGGCTATGTCGCGCTATTGGGTACACGCT	434				

This Page Blank (uspto)

181 CAGATGCTGATGACACCCGAGAGGCAATTTGAAGTTCCACAAAGAGCTCTGGGAAGTG 240
181 CAGATGCTGATGACACCCGAGAGGCAATTTGAAGTTCCACAAAGAGCTCTGGGAAGTG 240
241 CTGCTGATGACAGTTTGGCAACCACTCAACCTGTGGAAGGTGACTATTTTGGCTCGAG 300
241 CTGCTGATGACAGTTTGGCAACCACTCAACCTGTGGAAGGTGACTATTTTGGCTCGAG 300
301 TTTCCGATCAGAAAAGATCAACGGTGGCTGGATCTCTAAACCCCATTTGGAACAG 360
301 TTTCCGATCAGAAAAGATCAACGGTGGCTGGATCTCTAAACCCCATTTGGAACAG 360
361 ATTAGAAGCCAAAGACAGTTGTTGTTAAAGTTGTGTGAAATTCCTTCCGCTGACAC 420
361 ATTAGAAGCCAAAGACAGTTGTTGTTAAAGTTGTGTGAAATTCCTTCCGCTGACAC 420
421 ACACAACTCCAAAGAACTCAACAGGTACCTGTTCGCGCTGAGAGTGAAGACAGCTTG 480
421 ACACAACTCCAAAGAACTCAACAGGTACCTGTTCGCGCTGAGAGTGAAGACAGCTTG 480
481 GGTCAAGGAGGTTGACGTGTAATGACACAGGCGAGCTCTGTGATTTCAACAATTGG 540
481 GGTCAAGGAGGTTGACGTGTAATGACACAGGCGAGCTCTGTGATTTCAACAATTGG 540
541 CAATCTGAGATTGGGATTTTGTATGAAAGCTTGTGACAGAGACACTTAAACAAAATAA 600
541 CAATCTGAGATTGGGATTTTGTATGAAAGCTTGTGACAGAGACACTTAAACAAAATAA 600
601 TACATTACTCAGCAAGACGCACTAAGAGCAAAATGTGTGAATTTCAACATTAACCAATT 660
601 TACATTACTCAGCAAGACGCACTAAGAGCAAAATGTGTGAATTTCAACATTAACCAATT 660
661 GGACAAACACGAGCAAGATCAGATTTCCAGCTCCTAAGAGTTGCCGCTGAGAGATG 720
661 GGACAAACACGAGCAAGATCAGATTTCCAGCTCCTAAGAGTTGCCGCTGAGAGATG 720
721 TATGGAATCCGGTTGACCCGCGCAAGAGCAAGGAAAGCAAGAAATCAATCTGGCCGTT 780
721 TATGGAATCCGGTTGACCCGCGCAAGAGCAAGGAAAGCAAGAAATCAATCTGGCCGTT 780
781 GCCAAACGCGGAATTTAGTGTTCAGGGTTTCACTAAGATCAATGCTTCAACTGGGCC 840
781 GCCAAACGCGGAATTTAGTGTTCAGGGTTTCACTAAGATCAATGCTTCAACTGGGCC 840
841 AAGGTGCGGAAGCTGAGTTCAAGAGAGCGCTTTCATCAAGGTCGCGCAGATGTC 900
841 AAGGTGCGGAAGCTGAGTTCAAGAGAGCGCTTTCATCAAGGTCGCGCAGATGTC 900
901 AATAGTGCCTACACAGATACCTTGAATTCCTGATGGCCAGTCGGGATTTCTGCAAGTCC 960
901 AATAGTGCCTACACAGATACCTTGAATTCCTGATGGCCAGTCGGGATTTCTGCAAGTCC 960
961 TTTCTGAAAAATCTGTGTGAACATCATGCTTCTTGAAGCTTTTGAAGAGCCCAACCA 1020
961 TTTCTGAAAAATCTGTGTGAACATCATGCTTCTTGAAGCTTTTGAAGAGCCCAACCA 1020
1021 AAGCCCAAGCCGCTCTTTAGCCGCGGGATCATCTTCCGTTCAAGTGTGAGTCCAG 1080
1021 AAGCCCAAGCCGCTCTTTAGCCGCGGGATCATCTTCCGTTCAAGTGTGAGTCCAG 1080
1081 AAGCAGGTTCTGACTATGTAAAGAGAGACATTAAGAGGTGCACTTTGAAGAGAG 1140
1081 AAGCAGGTTCTGACTATGTAAAGAGAGACATTAAGAGGTGCACTTTGAAGAGAG 1140
1141 CACACAGAGATTCATTTCTATCCGGAAGCTTGTTCACAGCTTACAGAACTGAATTCGAA 1200
1141 CACACAGAGATTCATTTCTATCCGGAAGCTTGTTCACAGCTTACAGAACTGAATTCGAA 1200
1201 GTGCTGAGAGCTCTCAGAGAGACACAGCTTACATTTGAGAGAGGTGCGAATCTCCA 1260
1201 GTGCTGAGAGCTCTCAGAGAGACACAGCTTACATTTGAGAGAGGTGCGAATCTCCA 1260
1261 GGGGGCCAGAGCTGCGCGGAGAAAGAAACGAAGGTTTCGCGCGGGAGCCGGGGTTCG 1320

1261 GGGGGCCAGAGCTGCGCGGAGAAAGAAACGAAGGTTTCGCGCGGGAGCCGGGGTTCG 1320
1321 CACCCGAGCCCTGCGCCGAGAGAAAGCCCGGGTTAACAGACAGCGGACCGGCGC 1380
1321 CACCCGAGCCCTGCGCCGAGAGAAAGCCCGGGTTAACAGACAGCGGACCGGCGC 1380
1381 TCGGCGCCCAAGAGAGAGAGAGGTGCTTAAGATTAAGAACCCAGACAGTAACTT 1440
1381 TCGGCGCCCAAGAGAGAGAGAGGTGCTTAAGATTAAGAACCCAGACAGTAACTT 1440
1441 CAGCCCCGACGCAAGACAGAGCTCCCTGACTGAGCTCTCACTTTCCGAGCTGTCT 1500
1441 CAGCCCCGACGCAAGACAGAGCTCCCTGACTGAGCTCTCACTTTCCGAGCTGTCT 1500
1501 GTGAACCTGCAAGGGGGAGGTGCGCTTCCACACGTGACTTGTCTCCCAACTGAGCCCC 1560
1501 GTGAACCTGCAAGGGGGAGGTGCGCTTCCACACGTGACTTGTCTCCCAACTGAGCCCC 1560
1561 GACACCAAGCAGGCTCTCCCTTGATCAGCCGCTGCTGAATGAACAGGCTCTGCCCGG 1620
1561 GACACCAAGCAGGCTCTCCCTTGATCAGCCGCTGCTGAATGAACAGGCTCTGCCCGG 1620
1621 ACGAGATGAGATGAGGGCCGAGAGAGATTTCCCACTGATTAAGCGTACTTCATA 1680
1621 ACGAGATGAGATGAGGGCCGAGAGAGATTTCCCACTGATTAAGCGTACTTCATA 1680
1681 GCTAAGAGAGTCTACACACGAGAGCAATATCTGAAGATCTGGAAGTTATCACTTCG 1740
1681 GCTAAGAGAGTCTACACACGAGAGCAATATCTGAAGATCTGGAAGTTATCACTTCG 1740
1741 TGGTTTCAGACACAGTAGAGCAAGAGAGCCATGCGGAAAGCACTGAAAAGTCTCAT 1800
1741 TGGTTTCAGACACAGTAGAGCAAGAGAGCCATGCGGAAAGCACTGAAAAGTCTCAT 1800
1801 TTCCGAAATTTTGAACCTTTGCACAATTTCTACTAATTTTCTCAAGAAATTAAGCA 1860
1801 TTCCGAAATTTTGAACCTTTGCACAATTTCTACTAATTTTCTCAAGAAATTAAGCA 1860
1861 CGACTTGGCCCTGTGGGAAGGCGCTCAAAATGCCAAATCAGAGATTAACAAAGATCCGG 1920
1861 CGACTTGGCCCTGTGGGAAGGCGCTCAAAATGCCAAATCAGAGATTAACAAAGATCCGG 1920
1921 GATGTCATGCTGAAGAAATTCAGGGCATGAAGCACTGTGGGCTCACTGTGGAAGAC 1980
1921 GATGTCATGCTGAAGAAATTCAGGGCATGAAGCACTGTGGGCTCACTGTGGAAGAC 1980
1981 AGCGAGCCTTGGAGGCCCTGAGAGATGAATCAAGAGCTCCCGCGCTGGAAGACTTC 2040
1981 AGCGAGCCTTGGAGGCCCTGAGAGATGAATCAAGAGCTCCCGCGCTGGAAGACTTC 2040
2041 TGCAGAGCTTTGAGCTGACAGAGGTGTGTTACCTACAGGCTCAACACTTCTCTGGGG 2100
2041 TGCAGAGCTTTGAGCTGACAGAGGTGTGTTACCTACAGGCTCAACACTTCTCTGGGG 2100
2101 CCACTGCAACCGGCTCATGCACTACAGAGGTCCTGAGCGGCTGTGCAACACCAACCG 2160
2101 CCACTGCAACCGGCTCATGCACTACAGAGGTCCTGAGCGGCTGTGCAACACCAACCG 2160
2161 CCGAGCCACGCGGACTTCAAGGACTGCGGAGCGCTTTGGCAGAGATCAAGAGATGTTG 2220
2161 CCGAGCCACGCGGACTTCAAGGACTGCGGAGCGCTTTGGCAGAGATCAAGAGATGTTG 2220
2221 GCAAGCTCCACGGATGATGATCAAGATGAAGATTTCAAGAGCTGCAAGAACTCAAG 2280
2221 GCAAGCTCCACGGATGATGATCAAGATGAAGATTTCAAGAGCTGCAAGAACTCAAG 2280
2281 AAAGATTTGATTTGATGATCAATCTTGTGTTCCGGAAGAGGATTCATCCGCTGGGC 2340
2281 AAAGATTTGATTTGATGATCAATCTTGTGTTCCGGAAGAGGATTCATCCGCTGGGC 2340
2341 AGCTCAGCAAGCTCTCGGGGAAGGGGCTCAAGACGCAATGTTTCTCTGTTCAAGAC 2400

2341 AGCCTCAGAAAGCTCTCGGGGAAAGGGCTCAGAGGCACTTCTTCTTCTTCAAGAC 2400
QY 2401 GTCTCTGTATACAGAGCCGGGGGCTGACGGCTCCATATAGTTTAAAGTCCAGGGCAG 2460
DB 2401 GTCTCTGTATACAGAGCCGGGGGCTGACGGCTCCATATAGTTTAAAGTCCAGGGCAG 2460
QY 2461 CTCGCGCTCATATGAGCATGAGATTGAGGAGGAAAGACGAGTGGGGGGGCCCCACTGC 2520
DB 2461 CTCGCGCTCATATGAGCATGAGATTGAGGAGGAAAGACGAGTGGGGGGGCCCCACTGC 2520
QY 2521 CTGACCTCCGGGGCCAGCGCGAGTCCATCATGTGGCCGCACTTCTCGGTCCGAGATG 2580
DB 2521 CTGACCTCCGGGGCCAGCGCGAGTCCATCATGTGGCCGCACTTCTCGGTCCGAGATG 2580
QY 2581 GAGAAATGGGTTGAGGACATCCAGATGGCCATTGACTGGCGGAGAAAGACAGACAGCCC 2640
DB 2581 GAGAAATGGGTTGAGGACATCCAGATGGCCATTGACTGGCGGAGAAAGACAGACAGCCC 2640
QY 2641 GCGCCCTGAGTTCCTGGCCAGCAGCCGCCCTGACAAAGTCCGCTGATGAGGACAGCGG 2700
DB 2641 GCGCCCTGAGTTCCTGGCCAGCAGCCGCCCTGACAAAGTCCGCTGATGAGGACAGCGG 2700
QY 2701 GCTGACAGGAGTCAAGAGATGACTGAGCGCTCGCGACATCGCTGAGAGCGCAGGCGC 2760
DB 2701 GCTGACAGGAGTCAAGAGATGACTGAGCGCTCGCGACATCGCTGAGAGCGCAGGCGC 2760
QY 2761 CCGCACCGGGGCAACAAATGTGCAAGTGTGCTGGCAACCGCAACCGAGCTTCCATG 2820
DB 2761 CCGCACCGGGGCAACAAATGTGCAAGTGTGCTGGCAACCGCAACCGAGCTTCCATG 2820
QY 2821 GTGGAATTCAGCATCGCAGTGAAGAAATCACTTGTGGAAGAACTGCTGAAGAAATTC 2880
DB 2821 GTGGAATTCAGCATCGCAGTGAAGAAATCACTTGTGGAAGAACTGCTGAAGAAATTC 2880
QY 2881 AACAGCAAGGGGTGAGAGAGCTGTGGTGTGTGCAAACTTCTGCTGTCTTCTAC 2940
DB 2881 AACAGCAAGGGGTGAGAGAGCTGTGGTGTGTGCAAACTTCTGCTGTCTTCTAC 2940
QY 2941 AATCAACAGGAGCAATCATCCCTTGGCAGCTGTGCTGTGCTGCTGCTGCTGCTGCTG 3000
DB 2941 AATCAACAGGAGCAATCATCCCTTGGCAGCTGTGCTGTGCTGCTGCTGCTGCTGCTG 3000
QY 3001 ATCCCTCTGAGTCCGAGAAATCACTGAGTGTGCAAGTGTGCAAGTGTGCAAGTGTG 3060
DB 3001 ATCCCTCTGAGTCCGAGAAATCACTGAGTGTGCAAGTGTGCAAGTGTGCAAGTGTG 3060
QY 3061 CAGCTTACTACTTCAAGGCGGAAAGCAGTACAGCTTCCGAAAGTGTGATGAGATGATC 3120
DB 3061 CAGCTTACTACTTCAAGGCGGAAAGCAGTACAGCTTCCGAAAGTGTGATGAGATGATC 3120
QY 3121 CCGAGTGCACAGCTCTGCTGCGAAGCCCAAGTGTGAGCAAAAGATCTCTTGTG 3180
DB 3121 CCGAGTGCACAGCTCTGCTGCGAAGCCCAAGTGTGAGCAAAAGATCTCTTGTG 3180
QY 3181 TATTGATGCGGAGCACTCGTTCGAGAGTGGCTCTTCTGGAAGAGCTTCTCTT 3240
DB 3181 TATTGATGCGGAGCACTCGTTCGAGAGTGGCTCTTCTGGAAGAGCTTCTCTT 3240
QY 3241 CTTCTGTATTAATGAAGCTGTGTAATTAACACTGTCTGAAATCAAAAATCATGGCTT 3300
DB 3241 CTTCTGTATTAATGAAGCTGTGTAATTAACACTGTCTGAAATCAAAAATCATGGCTT 3300
QY 3301 CCGAGAGCTCTCTGCTGCTGCAAGCGGCTTTTAAACCCGAGCTCTCAGCGTTTGA 3360
DB 3301 CCGAGAGCTCTCTGCTGCTGCAAGCGGCTTTTAAACCCGAGCTCTCAGCGTTTGA 3360
QY 3361 TGAACAGCGCTCCACCTCAAGTCTGGCATCCGCTGGGGGGGCTGTTCTTTAGTATG 3420
DB 3361 TGAACAGCGCTCCACCTCAAGTCTGGCATCCGCTGGGGGGGCTGTTCTTTAGTATG 3420
QY 3421 CCAGTATTAACATTTGCTATT 3442
DB 3421 CCAGTATTAACATTTGCTATT 3442

RESULT 2
US-10-921-707-18
Sequence 18, Application US/10921707
Publication No. US20050003447A1
GENERAL INFORMATION:
APPLICANT: INCYTE PHARMACEUTICALS, INC.
APPLICANT: LAL, Preeti
APPLICANT: TANG, Y. Tom
APPLICANT: YUE, Henry
APPLICANT: HILLMAN, Jennifer L.
APPLICANT: BANDMAN, Olga
APPLICANT: CORLEY, Neil C.
APPLICANT: GUEGLER, Karl J.
APPLICANT: PATTERSON, Chandra
APPLICANT: AZIMZAI, Yalda
APPLICANT: BAUGHN, Mariah R.
TITLE OF INVENTION: HUMAN CYTOSKELETON ASSOCIATED PROTEINS
FILE REFERENCE: PF-0594 PCT
CURRENT APPLICATION NUMBER: US/10/921,707
CURRENT FILING DATE: 2004-08-19
PRIOR APPLICATION NUMBER: US/09/786,797
PRIOR FILING DATE: 2001-06-25
PRIOR APPLICATION NUMBER: 09/156,470; unassigned; 60/131,321
PRIOR FILING DATE: 1998-09-18; 1998-09-18; 1999-04-27
NUMBER OF SEQ ID NOS: 32
SOFTWARE: PERL Program
SEQ ID NO 18
LENGTH: 4687
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc. feature
OTHER INFORMATION: Incyte ID No: 1320252
US-10-921-707-18
Query Match 99.9%; Score 3437.2; DB 9; Length 4687;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3439; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 CCGCGCAGCGCGCGCGCGCTGTGAGATATTTCTTAAGCCGCTTTCATCATGAGAGAAATA 60
DB 42 CCGCGCAGCGCGCGCGCGCTGTGAGATATTTCTTAAGCCGCTTTCATCATGAGAGAAATA 101
QY 61 GAGCAGAGGCGCAGCCCAAGATCAAGCTGTGGGGCCCCGGAATTCGGGATCAAGTACC 120
DB 102 GAGCAGAGGCGCAGCCCAAGATCAAGCTGTGGGGCCCCGGAATTCGGGATCAAGTACC 161
QY 121 TTGGAACGTGAGCAGAAAGCGCGCCCAACACCTTCAGAGAAAATCTGTGTCTCAAAAATC 180
DB 162 TTGGAACGTGAGCAGAAAGCGCGCCCAACACCTTCAGAGAAAATCTGTGTCTCAAAAATC 221
QY 181 CAGATGCTGATGACACCCAGAGAGGATTTGAAGTTTCCCAAAAGAGCTCTCGGAAAGTG 240
DB 222 CAGATGCTGATGACACCCAGAGAGGATTTGAAGTTTCCCAAAAGAGCTCTCGGAAAGTG 281
QY 241 CTGTGTGATGAGTTTGCAACCACTCAACTGTGGAAGGTACTATTTTGGCTCGAG 300
DB 282 CTGTGTGATGAGTTTGCAACCACTCAACTGTGGAAGGTACTATTTTGGCTCGAG 341
QY 301 TTTCTGTATCAAAAAGATCAAGTGTGGCTGTGATTTCTTAAACCCATTTTGAAACAG 360
DB 342 TTTCTGTATCAAAAAGATCAAGTGTGGCTGTGATTTCTTAAACCCATTTTGAAACAG 401
QY 361 ATTGAAGGCAAAAGCAAGTGTGTAAGTTTGTGTAATTTCTTCCGCTGACAC 420
DB 402 ATTGAAGGCAAAAGCAAGTGTGTAAGTTTGTGTAATTTCTTCCGCTGACAC 461
QY 421 ACACAATCTCAAGAAAGATCAAGAGTACTGTTCGCTGAGAGTGAAGAGACATTTG 480
DB 462 ACACAATCTCAAGAAAGATCAAGAGTACTGTTCGCTGAGAGTGAAGAGACATTTG 521

Db	1602	GACACCAAGACGGCTCTCCCTTGATAGCCCGCTGTAATGACAGGCTTGC	1661
Qy	1621	ACGCACGATGAGGATGAGGGCCGGAGGAGATTTCCCACTGATAAAGCTACTTCA	1680
Db	1662	ACGCACGATGAGGATGAGGGCCGGAGGAGATTTCCCACTGATAAAGCTACTTCA	1721
Qy	1681	GCTAAGGAAGTGTCTACACCGAGCGAAACATCTGAAGGATCTGAAGTATCACTTCG	1740
Db	1722	GCTAAGGAAGTGTCTACACCGAGCGAAACATCTGAAGGATCTGAAGTATCACTTCG	1781
Qy	1741	TGTTTTCAGAGCACTGTAGCAAAAGAGACGCCATGCGGAGAGCACTGAAAGTCTATA	1800
Db	1782	TGTTTTCAGAGCACTGTAGCAAAAGAGAGCGCCATGCGCGAGACCTGAAAGTCTATA	1841
Qy	1801	TTCCCGAATTTTGAACTTTTGCAAAATTTCTATCTAATTTTCTCAAGGAAATTGAGCA	1860
Db	1842	TTCCCGAATTTTGAACTTTTGCAAAATTTCTATCTAATTTTCTCAAGGAAATTGAGCA	1901
Qy	1861	CGACTTCCCTGTGGAGAGGCCGCTCAATATGCCCAATCAGAGATTACAAAGATCGGC	1920
Db	1902	CGACTTCCCTGTGGAGAGGCCGCTCAATATGCCCAATCAGAGATTACAAAGATCGGC	1961
Qy	1921	GATGTCAATGCTGAAGAACTATTAGGCGATAGACACTTGCGCGCTCACTGTGAGACAC	1980
Db	1962	GATGTCAATGCTGAAGAACTATTAGGCGCATAGACACTTGCGCGCTCACTGTGAGACAC	2021
Qy	1981	AGCAGGACCTTGGAGGCGCTGGAGAAATGAATCAAGAGCTCCGCGCGCTGGAGAACTTC	2040
Db	2022	AGCAGGACCTTGGAGGCGCTGGAGAAATGAATCAAGAGCTCCGCGCGCTGGAGAACTTC	2081
Qy	2041	TGCAGAGACTTTGAGCTGACGAAGGTGTATTACCTACCGCTCAACACTTCTCTGCGG	2100
Db	2082	TGCAGAGACTTTGAGCTGACGAAGGTGTATTACCTACCGCTCAACACTTCTCTGCGG	2141
Qy	2101	CCACTGACACCGGCTCAATGCACTCAAGACAGCTCTGGAAGGCGCTGTCAAAACAACCGG	2160
Db	2142	CCACTGACACCGGCTCAATGCACTCAAGACAGCTCTGGAAGGCGCTGTCAAAACAACCGG	2201
Qy	2161	CCGAGCACAACCGCACTTCAAGGACCTGCGCAGACCGCTTTGGCAGAGATCAACGAGATGGT	2220
Db	2202	CCGAGCACAACCGCACTTCAAGGACCTGCGCAGACCGCTTTGGCAGAGATCAACGAGATGGT	2261
Qy	2221	GCACAGCTTCAACGGTACGATGATCAAGATGAGAGAAATTTCCAGAAAGCTGCAAGAACTCAAG	2280
Db	2262	GCACAGCTTCAACGGTACGATGATCAAGATGAGAGAAATTTCCAGAAAGCTGCAAGAACTCAAG	2321
Qy	2281	AAAGATTGATGGCAATTGACAAATCTTGTGATTCGGGAAAGGAGTTCATCCGCTGGGC	2340
Db	2322	AAAGATTGATGGCAATTGACAAATCTTGTGATTCGGGAAAGGAGTTCATCCGCTGGGC	2381
Qy	2341	AGCCTCAGCAAGCTCTTCGGGAAAGGGCTCCAGACGGCATGTTCTTCTGTTCACGCAC	2400
Db	2382	AGCCTCAGCAAGCTCTTCGGGAAAGGGCTCCAGACGGCATGTTCTTCTGTTCACGCAC	2441
Qy	2401	GTCCTGTATACACGAGCGCGGGGCTGACGGCTCCCAATCAGTTTAAAGTCAAGGCGAG	2460
Db	2442	GTCCTGTATACACGAGCGCGGGGCTGACGGCTCCCAATCAGTTTAAAGTCAAGGCGAG	2501
Qy	2461	CTCCCGCTTATGCAATGACGATTTGAGAGAGGAGAACGAGTGGGGGATGCCCATCTGC	2520
Db	2502	CTCCCGCTTATGCAATGACGATTTGAGAGAGGAGAACGAGTGGGGGATGCCCATCTGC	2561
Qy	2521	CTGACCTCTCGGGGCGACGCGCGACGTCAATATGTTGGCGCGCACTTTCGGTCCGAAGTG	2580
Db	2562	CTGACCTCTCGGGGCGACGCGCGACGTCAATATGTTGGCGCGCACTTTCGGTCCGAAGTG	2621
Qy	2581	GAGAAAGTGGTTGAGACATCCAGATGGCAATTGACCTTGGCGGAGAGAGAGACAGCCGC	2640
Db	2622	GAGAAAGTGGTTGAGACATCCAGATGGCAATTGACCTTGGCGGAGAGAGAGACAGCCGC	2681
Qy	2641	GCCCTGAGTTCTGTGCGACAGACCCCTTGACAAAGTCCCTGTATGAAGCAACGGCG	2700


```
Db 2682 GCCCTGAGTTCCTGGCCAGACGCCCCCTGACAAAGTCCCTGTATGAAGCCACCGCG 2741
Qy 2701 GCTGACCAAGAGTCAGAGATGACTTGAAGGCTCCGGCACTTGGCTGGAGCCGACAGGCC 2760
Db 2742 GCTGACCAAGAGTCAGAGATGACTTGAAGGCTCCGGCACTTGGCTGGAGCCGACAGGCC 2801
Qy 2761 CCGCAGCGCGGCAACAAATGTCAGCTGTGTCGACCGCAACACAGGCTCTCCATG 2820
Db 2802 CCGCAGCGCGGCAACAAATGTCAGCTGTGTCGACCGCAACACAGGCTCTCCATG 2861
Qy 2821 GTGAGCTTCAGATCCGAGTGGAGAAATCAATTGTCTGAAAACCTGTGAGAAATTCAAA 2880
Db 2862 GTGAGCTTCAGATCCGAGTGGAGAAATCAATTGTCTGAAAACCTGTGAGAAATTCAAA 2921
Qy 2881 AACAGCAACGGGTGGCGAAGAGCTGTGGTGTGTTCAAAATTTCTGCTTTCTTAC 2940
Db 2922 AACAGCAACGGGTGGCGAAGAGCTGTGGTGTGTTCAAAATTTCTGCTTTCTTAC 2981
Qy 2941 AAATCACACCAAGGACAAATCATCCCTTGCCAGGCTGCTGCTGGCTAATTGCTCACC 3000
Db 2982 AAATCACACCAAGGACAAATCATCCCTTGCCAGGCTGCTGCTGGCTAATTGCTCACC 3041
Qy 3001 ATCCCTCTGAGTCCGAGAACTTCAGAAAAGCTAGTGTTCAGAGCTGCACTTCAAGTCC 3060
Db 3042 ATCCCTCTGAGTCCGAGAACTTCAGAAAAGCTAGTGTTCAGAGCTGCACTTCAAGTCC 3101
Qy 3061 CAGCTCTACTACTTTCAGGGCGGAAAGCGAGTACAGTTCGAAAAGTGGATGAAAGTATC 3120
Db 3102 CAGCTCTACTACTTTCAGGGCGGAAAGCGAGTACAGTTCGAAAAGTGGATGAAAGTATC 3161
Qy 3121 CGCAGTGGCCACAGCTCTGCTGCGGCAACCCAGTGTGAGCCACAAAGAGTCTCTTG 3180
Db 3162 CGCAGTGGCCACAGCTCTGCTGCGGCAACCCAGTGTGAGTCAAAAGAGTCTCTTG 3221
Qy 3181 TATGATGGCCGAGCACTGTTTCGAGTGGCTGCTTCTGGAAGAGCTTCTT 3240
Db 3222 TATGATGGCCGAGCACTGTTTCGAGTGGCTGCTTCTGGAAGAGCTTCTT 3281
Qy 3241 CTTCGTATTTAATGAAGCTGTGTTAAATTAACCTGTCTGAAAATCAAAATCATGCTT 3300
Db 3282 CTTCGTATTTAATGAAGCTGTGTTAAATTAACCTGTCTGAAAATCAAAATCATGCTT 3341
Qy 3301 CCCAGAGCTCTCTGCTCTCCAGAGCGCGTTTTTTAAACCCGAGCTCTGAGGTTTTGAA 3360
Db 3342 CCCAGAGCTCTCTGCTCTCCAGAGCGCGTTTTTTAAACCCGAGCTCTGAGGTTTTGAA 3401
Qy 3361 TGAACAGCGCTCCCACTCCAGTCTTGATCCGCTGGGGGCGCTGTTTGTAGTAGTG 3420
Db 3402 TGAACAGCGCTCCCACTCCAGTCTTGATCCGCTGGGGGCGCTGTTTGTAGTAGTG 3461
Qy 3421 CCAGTATTTAAACATTTGCTATT 3442
Db 3462 CCAGTATTTAAACATTTGCTATT 3483

RESULT 3
US-10-450-763-564
/ Sequence 564, Application US/10450763
/ Publication No. US20050196754A1
/ GENERAL INFORMATION:
/ APPLICANT: Hyeeg, Inc
/ TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
/ FILE REFERENCE: 790CIPJ/US
/ CURRENT APPLICATION NUMBER: US/10/450, 763
/ CURRENT FILING DATE: 2003-06-11
/ PRIOR APPLICATION NUMBER: PCT/US01/08631
/ PRIOR FILING DATE: 2001-03-30
/ PRIOR APPLICATION NUMBER: 09/540,217
/ PRIOR FILING DATE: 2000-03-31
/ PRIOR APPLICATION NUMBER: 09/649,167
/ PRIOR FILING DATE: 2000-08-23
/ NUMBER OF SEQ ID NOS: 60736
/ SOFTWARE: Custom
```

```
/ SEQ ID NO 564
/ LENGTH: 3187
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: SIMILAR
/ LOCATION: (49)..(1371)
/ OTHER INFORMATION: 92% homologous to Homo sapiens CDP, accession number
/ OTHER INFORMATION: AB008430, Smith-Waterman Score=2090.
US-10-450-763-564

Query Match 92.1%; Score 3170.2; DB 10; Length 3187;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3183; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

Qy 1 CGCCGAGCGCGCGCGCTGTGAGATATTCTTAAGCCGCTTTCATCATGGAGAAATG 60
Db 1 CGCCGAGCGCGCGCGCTGTGAGATATTCTTAAGCCGCTTTCATCATGGAGAAATG 60
Qy 61 GAGCAGAGGCCAGCCCGAGATCAAGCTGGGGGCCCCGGAAAATTCGGAGATCACTACC 120
Db 61 GAGCAGAGGCCAGCCCGAGATCAAGCTGGGGGCCCCGGAAAATTCGGAGATCACTACC 120
Qy 121 TTGGAACGTGACAGAAAGCGCCGCCCAACCTTCAGAAAATCTGCTGTCATCAAAATC 180
Db 121 TTGGAACGTGACAGAAAGCGCCGCCCAACCTTCAGAAAATCTGCTGTCATCAAAATC 180
Qy 181 CAGATCTGATGACACCCAGAGGCAATTTGAATTCACAAAGACTCTCTGGAAAGTG 240
Db 181 CAGATCTGATGACACCCAGAGGCAATTTGAATTCACAAAGACTCTCTGGAAAGTG 240
Qy 241 CTGCTGGAGTGCATTTGCAACCACTCACTCTGAGAAAGTGAATTTGGCCCTCGAG 300
Db 241 CTGCTGGAGTGCATTTGCAACCACTCACTCTGAGAAAGTGAATTTGGCCCTCGAG 300
Qy 301 TTTCTGATCAAAAAGATCAAGGTGCTGATCTCTTAAACCAATTTGAAGACAG 360
Db 301 TTTCTGATCAAAAAGATCAAGGTGCTGATCTCTTAAACCAATTTGAAGACAG 360
Qy 361 ATTGAAGGCCAAAGCACTGTTTAAAGTTGTGTAATTTCTTCCGCTGACAC 420
Db 361 ATTGAAGGCCAAAGCACTGTTTAAAGTTGTGTAATTTCTTCCGCTGACAC 420
Qy 421 AACAACTCCAAAGAAATCAAGGTACCTGTCGCTGAGTGAAGAGGAGCTTG 480
Db 421 AACAACTCCAAAGAAATCAAGGTACCTGTCGCTGAGTGAAGAGGAGCTTG 480
Qy 481 GCTCAAGGAGGTGACGTGTATGACACAGGAGCTCTGTTGATTTCAACATTTTG 540
Db 481 GCTCAAGGAGGTGACGTGTATGACACAGGAGCTCTGTTGATTTCAACATTTTG 540
Qy 541 CAATCTGAGATGGGATTTTATGAAAGCTTTGACAGAGCACTTACCAATAATAA 600
Db 541 CAATCTGAGATGGGATTTTATGAAAGCTTTGACAGAGCACTTACCAATAATAA 600
Qy 601 TACATCTCAGAGAGAGCACTAGAGGACAAATTCGTGGAATTTTACATAACCAT 660
Db 601 TACATCTCAGAGAGAGCACTAGAGGACAAATTCGTGGAATTTTACATAACCAT 660
Qy 661 GGAACAAACAGAGAAATCAATTTCCAGCTCTAGAGATTTGCCGTGCTAGAGATG 720
Db 661 GGAACAAACAGAGAAATCAATTTCCAGCTCTAGAGATTTGCCGTGCTAGAGATG 720
Qy 721 TATGAAATCCGTTGACCCCGGCAAGGACAGGAAAGGACAGAAATCAATCTGCGCTT 780
Db 721 TATGAAATCCGTTGACCCCGGCAAGGACAGGAAAGGACAGAAATCAATCTGCGCTT 780
Qy 781 GCCAACAGGGAATCTAGTGTTCAGGGTTTCATGAATGAATGCTTCAATGGGCTC 840
Db 781 GCCAACAGGGAATCTAGTGTTCAGGGTTTCATGAATGAATGCTTCAATGGGCTC 840
Qy 841 AAGTGGAGAGCTGCTTCAAGAGAGCGCTTTCATCAAGCTCCGCGCAGATGCC 900
Db 841 AAGTGGAGAGCTGCTTCAAGAGAGCGCTTTCATCAAGCTCCGCGCAGATGCC 900
```

Db 841 AAGGTCGGAAAGCTTGACCTTCAAGAGAAAGCGCTTCTCATCAAGCTCCGGCCAGATGCC 900
QY |||||
Db 901 AATPAGGCGTACAGAGATACCTTGGAATCCGTGATGGCCAGTCGGGATTTCTGCAATGCC 960
QY |||||
Db 901 AATPAGGCGTACAGAGATACCTTGGAATCCGTGATGGCCAGTCGGGATTTCTGCAATGCC 960
QY |||||
QY 961 TTCTGGAAAATCTGTGTGNAACATCATGCTTCTTTGACTTTTGTGAAAGGCCCAACCA 1020
Db 961 TTCTGGAAAATCTGTGTGNAACATCATGCTTCTTTGACTTTTGTGAAAGGCCCAACCA 1020
QY 1021 AAGCCCAAGCCGCTCTTTAGCCGGGGTCACTATTTGGTTCTAGTGGTGGACTCG 1080
Db 1021 AAGCCCAAGCCGCTCTTTAGCCGGGGTCACTATTTGGTTCTAGTGGTGGACTCG 1080
QY 1081 AAGCAGGTTCTGACTATGTTTAAAGAGGACATPAGAGGTTGAGTTTGAAGAGAA 1140
Db 1081 AAGCAGGTTCTGACTATGTTTAAAGAGGACATPAGAGGTTGAGTTTGAAGAGAA 1140
QY 1141 CACAGCAAGATTCTATTCGAGACCTTGCTTCAAGCTTACAGACTGAATTCGAA 1200
Db 1141 CACAGCAAGATTCTATTCGAGACCTTGCTTCAAGCTTACAGACTGAATTCGAA 1200
QY 1201 GTGCTGGAGAGTCTCAGCAGAGCACCGCTTACATTTGGAGA -AGTGCAGATCTCC 1259
Db 1201 GTGCTGGAGAGTCTCAGCAGAGCACCGCTTACATTTGGAGAAGTGCAGATTTCC 1260
QY 1260 AAGGGGCGAGGCTGCGCGGAGAAAGAACCGAAGTTTCCGCGGGAGCGGGGTC 1319
Db 1261 AAGGGGCGAGGCTGCGCGGAGAAAGAACCGAAGTTTCCGCGGGAGCGGGGTC 1320
QY 1320 GCACCCGAGCCCTGCGCCGAGAGAAAGCCCGGGGTAAACAAGCAGCGGAGCGGAGCCG 1379
Db 1321 GCACCCGAGCCCTGCGCCGAGAGAAAGCCCGGGGTAAACAAGCAGCGGAGCGGAGCCG 1380
QY 1380 CTCGGGCCCCCAGAGAAAGAGAGAGTCTGTTAAGATAGACCCAGCAGATTAACC 1439
Db 1381 CTCGGAGCCCAAGAGAAAGAGAGAGTCTGTTAAGATAGACCCAGCAGATTAACC 1440
QY 1440 TCAGCCCGCGAGCAGAGCAGGCTCCCTGACTGGCAGTCCCTACCTTTCGAGCTGTC 1499
Db 1441 TCAGCCCTGCGAGCCAGAGCAGGCTCCCTGACTGGCAGTCCCTACCTTTCGAGCTGTC 1500
QY 1500 TGTGAACTGCGAGGGGGAGTGGCCCTGCGCAAGTGAAGCTTGTCTCCAGCTGAGCC 1559
Db 1501 TGTGAACTGCGAGGGGGAGTGGCCCTGCGCAAGTGAAGCTTGTCTCCAGCTGAGCC 1560
QY 1560 CGACACCAAGCAGGCTCTTCTTGTATCAGCCCGCTGCTGAATGACCAAGGCTGCCCCG 1619
Db 1561 CGACACCAAGCAGGCTCTTCTTGTATCAGCCCGCTGCTGAATGACCAAGGCTGCCCCG 1620
QY 1620 GACGAGCAGTGAAGTGAAGGCGGAGGAGAGATTCCCACTGATPAAAGCGTACTTCA 1679
Db 1621 GACGAGCAGTGAAGTGAAGGCGGAGGAGAGATTCCCACTGATPAAAGCGTACTTCA 1680
QY 1680 AGCTPAAAGAGTGTCTACACCGAGCGAATATCTGAAGAGTCTGAAGTTATCACTTC 1739
Db 1681 AGCTPAAAGAGTGTCTACACCGAGCGAATATCTGAAGAGTCTGAAGTTATCACTTC 1740
QY 1740 GTGGTTTCAGAGCAGTGAAGCAAGAGAGCGCATGCGGAAAGCACTGAAAAGTTCAT 1799
Db 1741 GTGGTTTCAGAGCAGTGAAGCAAGAGAGCGCATGCGGAAAGCACTGAAAAGTTCAT 1800
QY 1800 AATCCGGAATTTGAACCTTTGCAAAATTTCTATCTAATTTTCTCAAGAAATTGAGA 1859
Db 1801 AATCCGGAATTTGAACCTTTGCAAAATTTCTATCTAATTTTCTCAAGAAATTGAGA 1860
QY 1860 ACGACTTGCCTGTGGAAAGCGCTCAATGCCCCAATCAGATTAACCAAGATCCG 1919
Db 1861 ACGACTTGCCTGTGGAAAGCGCTCAATGCCCCAATCAGATTAACCAAGATCCG 1920
QY 1920 CGATGTCACTGTGAAGAACATTCAAGGCGATGAAGCCTGGCGGCTCACTGTGGAAGCA 1979
Db 1921 CGATGTCACTGTGAAGAACATTCAAGGCGATGAAGCCTGGCGGCTCACTGTGGAAGCA 1980

QY 1980 CAGCAGGCTTTGAGGCCCTGAGAAATGAATCAAGAGCTCCGGCGGCTGGAATCTT 2039
Db 1981 CAGCAGGCTTTGAGGCCCTGAGAAATGAATCAAGAGCTCCGGCGGCTGGAATCTT 2040
QY 2040 CTGCAAGACTTTGAGCTGCAAGAGTGTATCTTACCGCTCAACACTTCTCTCGG 2099
Db 2041 CTGCAAGACTTTGAGCTGCAAGAGTGTATCTTACCGCTCAACACTTCTCTCGG 2100
QY 2100 GCCACTGCAACCGGCTCATGCACTAACAAGGTCTTGAAGCGGCTGTGCAACACACCC 2159
Db 2101 GCCACTGCAACCGGCTCATGCACTAACAAGGTCTTGAAGCGGCTGTGCAACACACCC 2160
QY 2160 GCCGAGCAGCGCGACTTCAAGGAGCTGCGAGCCGCTTTGGCAGAGATCAAGAGATG 2219
Db 2161 GCCGAGCAGCGCGACTTCAAGGAGCTGCGAGCCGCTTTGGCAGAGATCAAGAGATG 2220
QY 2220 GGCACAGCTCCACGGTACGATGATCAAGATGGAATTTCCAGAACTGACGAATCTCA 2279
Db 2221 GGCACAGCTCCACGGTACGATGATCAAGATGGAATTTCCAGAACTGACGAATCTCA 2280
QY 2280 GAAAGATTGATTTGGCATTTGAACAATCTGTGTTCCGGGAAAGGAGTTCAATCCGTGG 2339
Db 2281 GAAAGATTGATTTGGCATTTGAACAATCTGTGTTCCGGGAAAGGAGTTCAATCCGTGG 2340
QY 2340 CAGCCTCAGCAAGCTCTGCGGGAAAGGCTCCAGCAGCGCATGTTCTTCTGTTCAACGA 2399
Db 2341 CAGCCTCAGCAAGCTCTGCGGGAAAGGCTCCAGCAGCGCATGTTCTTCTGTTCAACGA 2400
QY 2400 GGTCTGTCTATACAGAGCCCGGGGCTGACGGCTTCAATCAATGTTTAAATGTCACGGGCA 2459
Db 2401 GGTCTGTCTATACAGAGCCCGGGGCTGACGGCTTCAATCAATGTTTAAATGTCACGGGCA 2460
QY 2460 GGTCCCGCTATGTCATGATGATGAGAGAGAGAGCGAGTGGGGGAGTCCCACTG 2519
Db 2461 GGTCCCGCTATGTCATGATGATGAGAGAGAGAGCGAGTGGGGGAGTCCCACTG 2520
QY 2520 CCTGACCTCCGGGCGCAGCGCAGTCCATCATCTGTGCGCGCATGTTCTCGTCCGAGAT 2579
Db 2521 CCTGACCTCCGGGCGCAGCGCAGTCCATCATCTGTGCGCGCATGTTCTCGTCCGAGAT 2580
QY 2580 GAGAAATGTGTTGAGACATTCAGATGCGCAATGACTGTGCGGAGAAAGAGCAGCC 2639
Db 2581 GAGAAATGTGTTGAGACATTCAGATGCGCAATGACTGTGCGGAGAAAGAGCAGCC 2640
QY 2640 GCGCCCTGAGTTCCTGGCCAGAGCCCTGCAACAAGTCCCTGATGAAGCCAGCC 2699
Db 2641 GCGCCCTGAGTTCCTGGCCAGAGCCCTGCAACAAGTCCCTGATGAAGCCAGCC 2700
QY 2700 GACTGACAGAGTCAAGAGTGAAGCTGAGCGCTGCGCAATGCTGAGAGCGCAGGC 2759
Db 2701 GACTGACAGAGTCAAGAGTGAAGCTGAGCGCTGCGCAATGCTGAGAGCGCAGGC 2760
QY 2760 CCGCAGCCGCGCAACAATGTGTGCACTGTGCTGCGCAACCAAGCTGTCTCAT 2819
Db 2761 CCGCAGCCGCGCAACAATGTGTGCACTGTGCTGCGCAACCAAGCTGTCTCAT 2820
QY 2820 GGTGACTTCAAGCATCGAGTGAAGATCAAGTGTCTGAAAACCTGTGAGAAAATTCA 2879
Db 2821 GGTGACTTCAAGCATCGAGTGAAGATCAAGTGTCTGAAAACCTGTGAGAAAATTCA 2880
QY 2880 AAACAGCAACGGGTGCGAAGACTGTGGGTGTTCAAAATTCTGCTGTCTTCTA 2939
Db 2881 AAACAGCAACGGGTGCGAAGACTGTGGGTGTTCAAAATTCTGCTGTCTTCTA 2940
QY 2940 CAATCAACAACAGCAATCATCCCTTTCAGAGCTGCTGTGCTGCTGCTGCTGCTGCTG 2999
Db 2941 CAATCAACAACAGCAATCATCCCTTTCAGAGCTGCTGTGCTGCTGCTGCTGCTGCTG 3000
QY 3000 CATCCCTCTGAGTCCGAGAACATCAAGAAAGCTAAGTGTCAAGCTGCACTTCAAGTC 3059
Db 3001 CATCCCTCTGAGTCCGAGAACATCAAGAAAGCTAAGTGTCAAGCTGCACTTCAAGTC 3060

QY 3060 CCACGTCTACTACTTCAGAGGCGGAAAGCGAGTACACGTTGAAAAGGTGATGATGAT 3119
DB 3061 CCACGTCTACTACTTCAGAGGCGGAAAGCGAGTACACGTTGAAAAGGTGATGATGAT 3120
QY 3120 CCGCAGTGGCCACGAGCTTGCCTTCGCGAGCCCACTGTGTAGCCACAAAAGTCTCTTGT 3179
DB 3121 CCGCAGTGGCCACGAGCTTGCCTTCGCGAGCCCACTGTGTAGCCACAAAAGTCTCTTGT 3180
QY 3180 GTATTGA 3186
DB 3181 GTATTGA 3187
RESULT 4
US-09-925-297-220
Sequence 220, Application US/09925297
Patent No. US20020081659A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
FILE REFERENCE: PA105
CURRENT APPLICATION NUMBER: US/09/925,297
PRIOR FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: PCT/US00/05989
PRIOR FILING DATE: 2000-03-08
PRIOR APPLICATION NUMBER: 60/124,270
NUMBER OF SEQ ID NOS: 928
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 220
LENGTH: 3094
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (2)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (4)
OTHER INFORMATION: n equals a,t,g, or c
US-09-925-297-220
Query Match 49.1%; Score 1688.6; DB 3; Length 3094;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1659; Conservative 2; Mismatches 1; Indels 1; Gaps 1;
QY 1740 GTGGTTTCAGAGCAGTGAAGCAAGAGAGAGCGCCATGCCGAGCACTGAAAAGTCTCAT 1799
DB 62 GTGGTTTCAGAGCAGTGAAGCAAGAGAGAGCGCCATGCCGAGCACTGAAAAGTCTCAT 121
QY 1800 ATTCCCGAATTTGAACTTTGCACAAAATTTCTACTAATTTTCTCAAGAAATTGAGCA 1859
DB 122 ATTCCTCGAATTTGAACTTTGCACAAAATTTCTACTAATTTTCTCAAGAAATTGAGCA 181
QY 1860 ACGACTTGCCTGTGGGAAGCGGCTCAAAATGCCCAATCAGAGATTACCAAAAGATCGG 1919
DB 182 ACGACTTGCCTGTGGGAAGCGGCTCAAAATGCCCAATCAGAGATTACCAAAAGATCGG 241
QY 1920 CGATGTCTATGCTGAGAAACAATTCAGGCGATGAAGCACTTGCAGGCTCACTGTGAAGCA 1979
DB 242 CGATGTCTATGCTGAGAAACAATTCAGGCGATGAAGCACTTGCAGGCTCACTGTGAAGCA 301
QY 1980 CAGCGAGGCTTGGAGGCGCTTGGAGAAATGGAATCAAGAGTCCCGCGGCTGAGAACTT 2039
DB 302 CAGCGAGGCTTGGAGGCGCTTGGAGAAATGGAATCAAGAGTCCCGCGGCTGAGAACTT 361
QY 2040 CTSCAGAGACTTTGAGCTGCAAGAGGTGTGTACTTACCGCTCAACACCTTCTCTGCG 2099
DB 362 CTSCAGAGACTTTGAGCTGCAAGAGGTGTGTACTTACCGCTCAACACCTTCTCTGCG 421
QY 2100 GCCACTGCACCGGCTCATGCACTACAGACAGGTCTTGAAGCGGCTGTGCAACCAACC 2159
DB 422 GCCACTGCACCGGCTCATGCACTACAGACAGGTCTTGAAGCGGCTGTGCAACCAACC 481

QY 2160 GCCAGGCCACGGCGAATTTGAGGGAATGCGGAGCGGCTTTGGAGAGATCAGCGAATGAT 2219
DB 482 GCCAGGCCACGGCGAATTTGAGGGAATGCGGAGCGGCTTTGGAGAGATCAGCGAATGAT 541
QY 2220 GGCAACACTTCAACGATGATGATCAAGATGAGAAATTTCCAGAAAGTGCAGAACTCAA 2279
DB 542 GGCAACACTTCAACGATGATGATCAAGATGAGAAATTTCCAGAAAGTGCAGAACTCAA 601
QY 2280 GAAAGATTTGATTGGCATTTGACATTTGTGTTCCGGGAAGGAGTTCACTCGTCTGGG 2339
DB 602 GAAAGATTTGATTGGCATTTGACATTTGTGTTCCGGGAAGGAGTTCACTCGTCTGGG 661
QY 2340 CAGCCTCAGCAAGCTCTCGGGGAAGGGGCTCCAGACGCGCATGTTCTCTGTCAACA 2399
DB 662 CAGCCTCAGCAAGCTCTCGGGGAAGGGGCTCCAGACGCGCATGTTCTCTGTCAACA 721
QY 2400 CGTCTGCTATACAGAGCGGGGGCTGACGGCTCTCAATCAGTTTAAAGTCCACGGGCA 2459
DB 722 CGTCTGCTATACAGAGCGGGGGCTGACGGCTCTCAATCAGTTTAAAGTCCACGGGCA 781
QY 2460 GCTCCCGCTCTATGCGATGACATTTGAGAGACCGAAGCAGAGTGGGGGCTGCCACTG 2519
DB 782 GCTCCCGCTCTATGCGATGACATTTGAGAGACCGAAGCAGAGTGGGGGCTGCCACTG 841
QY 2520 CCTGACCCCTCGGGGCGACGGGCGAGTCCATCATGCGGGCGGCAAGTTCTCGGTCGAGAT 2579
DB 842 CCTGACCCCTCGGGGCGACGGGCGAGTCCATCATGCGGGCGGCAAGTTCTCGGTCGAGAT 901
QY 2580 GGAAGATGGGTTGAGAGCATCCAGATGCGCAATTCGACCTGGCGAGAAAGCAGACGCC 2639
DB 902 GGAAGATGGGTTGAGAGCATCCAGATGCGCAATTCGACCTGGCGAGAAAGCAGACGCC 961
QY 2640 CGCCCTGAGTTCTGCGCAGACAGCCCTCTGACAAAGTCCCTGATGAAGCAACGCC 2699
DB 962 CGCCCTGAGTTCTGCGCAGACAGCCCTCTGACAAAGTCCCTGATGAAGCAACGCC 1021
QY 2700 GGCTGACGAGGATGACAGAGATGACCTGAGGCGCTCGGCAATGCTGAGACGCGAGGC 2759
DB 1022 GGCTGACGAGGATGACAGAGATGACCTGAGGCGCTCGGCAATGCTGAGACGCGAGGC 1080
QY 2760 CCGGACCGCGGCAACACATGATGTCAGTGTGCTGAGCACCGCAACACCGAGCTCTCAT 2819
DB 1081 CCGGACCGCGGCAACACATGATGTCAGTGTGCTGAGCACCGCAACACCGAGCTCTCAT 1140
QY 2820 GGTGACCTTCAGATCGCAGTGAAGATCATTTGCTGAGAACTTGTGAGAAATTCAA 2879
DB 1141 GGTGACCTTCAGATCGCAGTGAAGATCATTTGCTGAGAACTTGTGAGAAATTCAA 1200
QY 2880 AAAAGCAACGGGTGGCAAGACTGTGGGTGGTTCAAACTTCTGCTGTTCTCTTA 2939
DB 1201 AAAAGCAACGGGTGGCAAGACTGTGGGTGGTTCAAACTTCTGCTGTTCTCTTA 1260
QY 2940 CAATATCACACAGACCAATCATCTCCCTTGCAGCCTGCTCTGCTCAGTACTCTGCTCAC 2999
DB 1261 CAATATCACACAGACCAATCATCTCCCTTGCAGCCTGCTCTGCTCAGTACTCTGCTCAC 1320
QY 3000 CATCCCTTGAATCGAGAACATTCAGAAAGACTACGTGTTCAAGCTGCACTTCAAGTC 3059
DB 1321 CATCCCTTGAATCGAGAACATTCAGAAAGACTACGTGTTCAAGCTGCACTTCAAGTC 1380
QY 3060 CCAAGCTTACTTCACTTCAAGGCGGAAAGCGAGTACAGTTGCAAGAGTGAATGGAAGTAT 3119
DB 1381 CCAAGCTTACTTCACTTCAAGGCGGAAAGCGAGTACAGTTGCAAGAGTGAATGGAAGTAT 1440
QY 3120 CCGCAGTGGCCACGAGCTTGCCTTCGCGAGCCCACTGTGTAGCCCAAAAGATCTCTTGT 3179
DB 1441 CCGCAGTGGCCACGAGCTTGCCTTCGCGAGCCCACTGTGTAGCCCAAAAGATCTCTTGT 1500
QY 3180 GTATTGATGGCGGACACACTGTTTCCGACAGTGGCTGTTCTGTGAAAGAGTTCCTT 3239
DB 1501 GTATTGATGGCGGACACACTGTTTCCGACAGTGGCTGTTCTGTGAAAGAGTTCCTT 1560

```
QY 3240 TCTTGTATTAATGAGCGCTGTAAATTAACACCTGTCTGTAATAATCAAAAAATCATGCT 3299
DB 1561 TCTTGTATTAATGAGCGCTGTAAATTAACACCTGTCTGTAATAATCAAAAAATCATGCT 1620
QY 3300 TCCAGAGCTCTCTGTCTCTCCAGAGCGCGGTTTTTAAACCCGACCTCTCAGCGTTTGA 3359
DB 1621 TCCAGAGCTCTCTGTCTCTCCAGAGCGCGGTTTTTAAACCCGACCTCTCAGCGTTTGA 1680
QY 3360 ATGAAAGAGGCTCCCACTCCAGTCTCGGATCCGCTGGGGGGGCGTTCTTTAGCTAGT 3419
DB 1681 ATGAAAGAGGCTCCCACTCCAGTCTCGGATCCGCTGGGGGGGCGTTCTTTAGCTAGT 1740
QY 3420 GCCAGTATTTAAACATTTGCTATT 3442
DB 1741 GCCAGTATTTAAACATTTGCTATT 1763

RESULT 5
US-10-172-118-1536
; Sequence 1536, Application US/10172118
; Publication No. US20030224374A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Lineley, Peter
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Chris
; APPLICANT: Van 't Veer, Laura
; APPLICANT: Bernarde, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-175-999
; CURRENT APPLICATION NUMBER: US/10/172,118
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/380,770
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 1536
; LENGTH: 3997
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: NM_014808
; DATABASE ENTRY DATE: 2001-06-18
US-10-172-118-1536

Query Match 30.1%; Score 1037.4; DB 7; Length 3997;
Best Local Similarity 60.2%; Pred. No. 1.1e-305;
Matches 1864; Conservative 0; Mismatches 1186; Indels 45; Gaps 7;

QY 69 GCCGACCCGAGATCAAGACTGGGGGCCCGGAAAAATTCGGGGATCAGTACCTTGGAAAG 128
DB 150 GCAGACTGAGGGAGATCGCTGGGTGCCAGACCCCTGTGGAGTTAGACCTTGAGGC 209
QY 129 TGGACGAAGCGCGCCCAACACCTTCAGGAAAATCTGTGTCCATCAAAAATCCAGATGCT 188
DB 210 TGGGCGAGCTCTCTTCCCAAGATGCAAGAAAGCACTGCACCTCAGATGAAGCTGCT 269
QY 189 GAGTGCACCCGAGAGGACTTTGAAGTTCACAAAAGACTCCTCGGGAAGTGTCTGTGA 248
DB 270 GGAACAACACCATGGAATATTGACATTAAGCTTAATGCGAGTGGCAGGTATTACTGAC 329
QY 249 TGCAGTTTCAACCACTCAACCTCGTGAAGGTGACTATTTTGGCTCGAGTTTCTGA 308
DB 330 ACAAGTGTGAAGCTTTAACTGTGATGATGCTACTTCCGGATGGAAGTTTCAAA 389
QY 309 TCACAAAAGATCAAGGTGTGTGCTGATCTCTAAACCCATTTGTAACAGATTAGAG 368
DB 390 TACTAGTCTTACTGATTTGGCTGAACCTATGAACCAATCATTAAGCAAAATAGAG 449
QY 369 GCCAAAGCAGTTTGTGTTAAGTTTGTGTGAATCTTTCGGCTGACACACCAACT 428
DB 450 GCCAAAGATGTGTGCTTGTGCTGCTAGCTGTAATTTTTCACCTGATCTGTGTAGCT 509
```

```
QY 429 CCAGAAAGACTCAGAGGTACTGTTCGCGCTGAGGTGAAGCAGACTTGGCTCAAG 488
DB 510 ACAGAAAGATTAACAGATATCTTGTTCCTTGCAACTTAAGAGACGCTGTGAAGA 569
QY 489 CAGTTGACGTGTAAATGACACAGCGCAGCTCTCTTGAATTCACACATTTGCAATCTGA 548
DB 570 GCGTTTGACGTGTGTGACACACAGCGGCTTCTCAGTCCCATCTCTCCACTGCGA 629
QY 549 GATTGGGATTTTGTATGAAGCTTGGACAGAGAGCACTTAGCAAAAATTAATCATACC 608
DB 630 AATAGAAATTAAGATGAAGCGTGGACCGAGAGCACTCAAGTGAACGAGATTTGGC 689
QY 609 TCAGCAAGACGACATAGAGCAAAATCGTGAATTTCCATTAACCAATTTGACAAAC 668
DB 690 TGGCAGACGACCTGCTTGAAGATATCTGAATTCATGAGAGACGCTGGGCGACAGC 749
QY 669 ACCAGCAATCAATGATTTTCAGCTCTTAGAGATTGCCGTGGCTAGAGATGTATGANT 728
DB 750 ACTCTGAGTCTGATTTTCAGGTGCTCGAAATTCGAAAGTTGAAATGTACGGCAT 809
QY 729 CCGGTGACACCGGCAAGGACAGGGAGGACAGAAATCAATCGGCGTTGGCCACAC 788
DB 810 CAGATTTGACATGTCTTCTGACAGGGAGAAACAAAGTTCAACTGSCAGTTTCCACAT 869
QY 789 GGGAAATCTAGTGTTCAGGGTTTCACTAATCAATCCTTCAACTGSGCCAGGTGCG 848
DB 870 GGGTACTCTGTTGTTCCAGGGGACACCAAAATCAACTTTCAACTGTTCCAGTCCG 929
QY 849 GAGCTGAGCTTCAAGAGAAAGCGTTTCTCATCAAGCTCCGCGCAATGCCAATATGTC 908
DB 930 TAAATTAAGCTTCAAGAGAAAGATTTCTTATCAAACTTATCAAGGTTCATGAGCC 989
QY 909 GTACCAAGATPCTTGAATTCCTGATGSCAGTGGGATTTCTGCAAGTCTTCTGGA 968
DB 990 TTACAGAGACATTAAGATTTTGTGGTATGAGATGAATTAAGAACTTCTGGA 1049
QY 969 AATCTGTGTAAACATCATGCTCTCTTGAATTTTGAAGGCCAAACCAAGCCCA 1028
DB 1050 GATTTGTGTGAGATTCACACCTTTTGAATTTTGAATTCGAATTCGAATTCGA 1109
QY 1029 GCCGCTCTTTCAGCGGGGCTCATCTTCGTTGCTGTGAGTGTGAGTCTCAGAGCAGT 1088
DB 1110 AGCCGCTCTTTCAGCGGGGCTCTCTTCAAGATCAAGTGAAGAACTCAGAAACA 1169
QY 1089 TCTGACTATGTAAAGAGGACATTAAGAGGTGAGTTTGAAGAGAAAGACAGCA 1148
DB 1170 AGTAAATTTTCAAGAGCAGTGAATGAAGAAATTCATATGAAGAAAGCAGCA 1229
QY 1149 GATTCATTTCTATCCGAGCTTGTCTTCAAGCTTACAGAACTGAATTCGAAGTCTGGA 1208
DB 1230 GACCCACAGTCC-----GTTGAGCTCTGACTGACAGACTTCAAA 1271
QY 1209 GAGTCTCAGAGACACACAGCTTCACTTTGAGAGAGAGTCCGAAATCTCCAGGGGCA 1268
DB 1272 ACAGAGCATCTCATTTCCCGAGGAGTTGAGAGCTTCTTCCATCTTTCAGCAATGC 1331
QY 1269 GAGCTGCGGCGAAGAGAAAGAACCGAAGTTTCCGCGGGAGCGGGGTGCAACCGAG 1328
DB 1332 CTTTACTTCTGCTCTTCTTCCATCTGTGCTCTTCTGCTGCGCAGAGTTTAAAGAG 1391
QY 1339 CCCTGCGCGAGAGAGAACCCCGCGGTTAACAGAGAGCGGACGAGGCGCTCTGCGGCC 1388
DB 1392 CAGCAGCTCCCTCAGAGATCCCGAGGTTTCTTCAAGTCAAGAGTCCAGCTGCAAGAGGCG 1451
QY 1389 CACGAGAGAAAGAGAGAGG---TGTTAAGATTAAGACCCAGCAGAGATTAACCTCAGCC 1445
DB 1452 CAGTGAACAGTGTGTGAGGCGCCGACACACATCGGCCCAAGCCCTCGGGCCCCCGC 1511
QY 1446 CCGGAGCAAGCAGAGCTCCCTGACTGCGAGTCTTCACTTTTCGAGTGTCTGTGA 1505
DB 1512 ACTCAGGCTGTCCAGGCTTTTCCAGAGAGTCTTCAAGCTTCTCTCTCCAGCGGAA 1571
```

1506 CTGCAAGGGGAGTGGCCCTGCTCAACGTAACCTTGTCTCCCACTGAGCCCGAC 1565
1572 GAGCCCCCTGAGTCTGAGCCCTGATTTTCAAGTGCCTT-----TGGGCCAAGCTGA 1622
1566 CAAGCAGGCTCTCCCTTGTATCAGCCGCTGTGAATGACAGAGCTTCCCGGACGGA 1625
1623 ACAGGGCTATCCCACTCTGAGCCCTGTCTCAAGTATGTGGGAGACCGGAGATGGA 1682
1626 CAGTATGATGAGGGGCGGAGAGAGATTTCCCACTGAATTAAGCGTACTTCTAGCTAA 1685
1683 CTGCGAG---GAGCCAGACACAAGGCGCTGTGCTGACAGAGAGCGCTTACTTCAATCTCA 1739
1686 GGAAGTGTCTACACCGAGCGAATATATCTGAAGATCTCGAAGTTATCATCTTGTGTT 1745
1740 AAGATTTCTGCTACAGAAAGAACTATCTCAAGATTTTGAAGTTTATTCGTGTGTT 1799
1746 TCAGACACAGTGAAGAAAGAGACCGCATGCGGAGACATGAAAAAGTCTATTTTCCC 1805
1800 CCGCAGCGAGTGTGAAGAGAGCGCATGCTGCGACTCTGATGACGCTGCTTCTTC 1859
1806 GAATTTTGAACCTTTGACAAATTTTCTACTAATTTTCTCAAGAAATTTAGCAAGACT 1865
1860 CAACATCGATCCCATCTATGAGATTCCACAGAGGCTTCTGCGAGAGTGAAGAGAGCT 1919
1866 TGCCCTGTGGAGAGGCGCTCAATGCGCAATCAGAG---ATTACCAAGATTCGAGCA 1922
1920 GGCATCTTGGAGAGGCGCTCCAAAGCCACACAAAGAGATCATCAAGATTCGAGGA 1979
1923 TGTGATGCTGAAGAACATTCAGGCGATGAAGACCTGGCGGCTCACCTGTGAAGACAG 1982
1980 CATCTGCTCAGAGACATGCGCAGATTAAAGAGTTTACAGCTACTTCAAAAGACATGA 2039
1983 CGAGGCTTGAAGCCCTGTGAAGATGAAGATCAAGAGCTCCCGCGCTGTGAAGACTTTCG 2042
2040 CAGAGTCTTAACGAAGCTGAAAAAGGCTACCAAAACCTGTAAAGATTTGAAGCAGATGA 2099
2043 CAGAGCTTGAAGCTCAAGAAAGGTGTGTTACTACGCTCAACACCTTCTCTGCGGCT 2102
2100 CAGAGATTTGAAGCTCAAGAAAGTCTGTCTTCTGCTTCAACAGCTTCTGTTGAAGCC 2159
2103 ACTGCAACGCGCTCATGCACTCAAGAGAGTCTGTAGAGCGCTGTGCAAAACCAACCGCT 2162
2160 CATCCAGCGGCTGCTCACTACGCGCTGTGCTGCGCCGCTGTGCGACATTAACAGCC 2219
2163 GAGCCAGCGGCTTCAAGGAGCTGCGAGCGCTTGTGCAAGATCAAGAGATGTGCTG 2222
2220 CCGGACCACTGATGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTG 2279
2223 ACAGCTCAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2282
2280 CACATCAACGACATTTCTCTCTGCTGAGAACTGTCAAGAACTTAAAGAGCTGCAAGCG 2339
2283 AGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTG 2342
2340 GAGCCTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2399
2343 CCTCAGCAGGCTCTGCGGGAAGGAGCTTCAAGCAGCGGATTTTCTTCTGTTCAACAGCT 2402
2400 CTTTCAAGGCTCAACAAAGGCGCTGACAGAGAGATGTTTTTCTGTTCTCAATAT 2459
2403 CTTGCTATACAGAGCGGAGGCTGACGCGCTCAATCAAGTTTAAAGTCAAGGAGGCT 2462
2460 GTTGTCTGACAGAGAAAGAGATTTGAGAGACCAAGCTTCCGATTCGAGGCGCTCTCT 2519
2463 CCGCTCTATGCGATGAGATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2522
2520 TCCCTCTCAAGGATCTGTGTGAGAAAGATGATTAAGAGTGTCTGTTCAACAGTGT 2579
2523 GACCTTCCGCGGACAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2582
2580 CACCATCTACGCGGCTCAAGAAACAAATCTGTGTGAGAGCAGCAGCAGCAGCAGCAG 2639
2583 GAAGTGGTTGAGAGATCATGATGCTGATGCTGCGAGAGAGAGAGAGAGAGAGAGAG 2642

2640 GAAATGATGCTGGAATCTGCGGATTCACAGAGCAGCAAGAGTGGCGGTGACAGCGC 2699
2643 CCTGATGCTCTGGCCAGCA-----GCCCTCTGCAACAAAGTCCCTGATGAGAGCAC 2696
2700 CCTGACAGTCCAGAGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATG 2759
2697 CCGGCTGACAGAGATGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2756
2760 TCTG---GAGCAGAGATCAAGAGATGATGATGATGATGATGATGATGATGATGATGAT 2816
2757 GGCCTCCGACCGCGGACACAGATGATGATGATGATGATGATGATGATGATGATGATG 2816
2817 TGCCACAGACCGGCGCAACACAGATGATGATGATGATGATGATGATGATGATGATGAT 2876
2817 CATGTGACTTACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2876
2877 CAGGCGAGACCAAGTGAAGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 2936
2877 CAAAAACAGCAACGCTGAG 2936
2937 CAAAAACAGTCAATGCTGAG 2996
2937 CTACAAATCAACACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 2996
2997 CTACAAATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3056
2997 CACATCCCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3056
3057 GAGATCCCGCAGGAGGCGGATGATGATGATGATGATGATGATGATGATGATGATGAT 3116
3057 GTCCACGCTCTACTCTTACAGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3116
3117 ATCCACGCT 3176
3117 GATCCGAGTGGCCACGAGCTGTGCTGCGAGCC 3151
3177 GATCCAGGCGGCGACAGAGCTCAAGCGGAGAGGCC 3211

RESULT 6
US-10-342-887-1536
; Sequence 1536, Application US/10342887
; Publication No. US20040058340A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter S.
; APPLICANT: Mao, Mao
; APPLICANT: Roberte, Christopher J.
; APPLICANT: Van de Veer, Laura Johanna
; APPLICANT: Van de Veer, Marc J.
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-188-999
; CURRENT APPLICATION NUMBER: US/10/342, 887
; PRIORITY FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: 60/298, 918
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/380, 710
; PRIOR FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: 10/172, 118
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 1536
; LENGTH: 3997
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-342-887-1536

Query Match 30.1%; Score 1037.4; DB 8; Length 3997;
Best Local Similarity 60.2%; Pred. No. 1.1e-305;
Matches 1864; Conservative 0; Mismatches 1186; Indels 45; Gaps 7;

69 GCGAGCCCGAGATCAGACTGGGGGCCCCGGAAAAATTCGGGGATCAGTACCTTGGAAAG 128
150 GCGAGCTGAGGGATGCGCTTGGGTGCCAGAGCCCCCTGGGGATTAGACCTTGAAGCC 209
129 TGGACGAAGCCGCCCCCAACACTTCAGAAAAATCGTGTCCATCAAAATCCAGATGCT 188
210 TGGGAGACTCTTCTTCCAGAAATGCAAGAAAGACCTGCACCTCAGATTAAGCTGCT 269
189 GGAATGACCCCGAGGAGCATTTGAATTGCCAAAGAGCTCCGGGGAAGGTGTGTGGA 248
270 GAGCAACACCATGGAATATTGACATGAGCTTAATGCGATGCGCAGGTATTACTGAC 329
249 TCGAGTTTCAACACACTCAACCTCGTGGAGGTGACTATTTTGGCTCGAGTTTCTGA 308
330 ACAAGTGTGGAGCGTTTAAACCTGGTAGAATGTGACTACTTGGGATGGAAGTTTCAAA 389
309 TCACAAAAAGATCACGGTGTGCTGGATCTCTAAACCCTATTGAAAACAGATTAGANG 368
390 TACTCAGTCTTACTGATTGTGGCTTGAACCTATGAAACCCATCATATTAGCCAAATACGAAG 449
369 GCGAAGACGTTGTGTTAAGTTTGGTGAATCTTTCGCGCTGACACACACACT 428
450 GCGAAGAAATGTGTCTTCTGCTAGCTGTAAATTTTTCACCTGATCTGTGTAGCT 509
429 CCAAGAGAATCTCACAAGGTACCTGTTCCGCTGAGGTGAAGCAGAGACTTGGCTCAAG 488
510 ACAAGAGAATATACAGAAATCTTGTGTGCTTGCATTAAGAGAGACTGTGGAAAG 569
489 CAGGTTCAGCTGTATGACACACAGCGAGCTCTCTTGAATTTCAACATTTGTGCAATCTGA 548
570 GGGTTGACCTGTGCTGACACACAGCGGCCCTTCTCAGTCCCATCTCTCAGTCCGA 629
549 GATTGGGATTTTGAAGAGCTTGGACAGAGAGCACTTAGCAAAAAATTAATATACATACC 608
630 AATAGAGATATAGATGAAACGCTGSAACGAGAGCACTCAAGTGAAGAGATTTGCTC 689
609 TCAGCAAGACGCACTAGAGACAAATTCGTGNAATTTCACTAACACATTTGSAACAAAC 668
690 TGGCCAGAGCACTGCTTGAAGAAATATCTAAGATTTCAATCAGAAACGATGGGCCAGAC 749
669 ACCAGCAAAATCAGATTTTCCAGCTCTTAGAGATGCCCCGTGGCTAGAGATGTATGAAT 728
750 ACTGTGTGTCGGAATTTCCAGGTGCTCCAAATTTGCTCCAAAGTTGGAATATGACCGCAT 809
729 CCGGTTGCAACCGGSCCAAGGACAGGGAGGCAAGAGATCAATTCGCGGTTGCCAACAC 788
810 CAGATTTCAATGCTTCTGACAGGGAGAGAACCAAGATTTCAACTGGCAGTTTCCACAT 869
789 GGGAAATTCATGTTTCAAGGTTTCACTAAGATCAATGCTTCAACTGGGCCAAGGTGCG 848
870 GGGGTGACTCGTGTTCAGGGGACACACAAATTCACATTTCAATGGTCCAAAGGTCCG 929
849 GAAAGTGAAGCTTCAGAGAGAGAGCGCTTTCATCAAGCTCCGGCCAGATGCCAATATGTC 908
930 TAAACTTAACCTTCAGAGAGAAAAAGATTTCTTATCAAACTTCATCCAGAGTTTCATGSAAC 989
909 GTACCAAGATTAACCTTGGAAATTCCTGATGGCAGTGGGGAATTTCTGCAATGCTTCTGSA 968
990 TTACCAAGGACATTAAGAAATTTTGTGGGTAGTAGAATGATTAAGATCTTCTGGAA 1049
969 AATCTGTGTGAACATCAGCTTCTTATGACTTTTGAAGGCCCAACCAAGCCCA 1028
1050 GAATTTGTGTGAAGTACACACTTTTATGACTTTTGAACCAACTTAACCAAGGCAAA 1109
1029 GCGCGTCTCTTTAGCGGGGGGTCACTTTGCGTTCAAGTGTGCGACTCAGAGAGAGGT 1088
1110 AGCCCGTCTTCTCAGCGGGGGCTCTCTCTTCAGATTCAGTGAAGAACTCAGAAACACT 1169
1089 TCTCGACTATGTTAAAGAGAGAGACATTAAGAGGCAAGTTTGAAGAGAGACAGCA 1148
1170 AGTAGTTATTTCAAGACAGTGAATGAAGAAATTCATATGAAGAGAGCAGCA 1229

1149 GATTCATTCTATCCGAGACTTGTCTTCAAGCTTACAGAACTGAATTCGAAAGTCTGGA 1208
1230 GACCCACACGTC-----GTTGAGCTGTGACTGACAGACTTACCAAA 1271
1209 GCAATCTCAGCAGAGACACAGCTTACATTTGGAAGAGGTGCGGAATCTCCAGGGGCCA 1268
1272 ACAGAGCATTCATATTCGCGAGGAGATTGAGGACTTCCTCCTCCCATCTTTCAGCCAAATGC 1331
1269 GAGCTGCCGCGAGGAAAGAAACGAAAGTTTCCGCGGGGAGCCGGGGTGCACCCGAG 1328
1332 CTTTAACTGCTCTCTCCCTCAGCTCTGTGCTCCCTCTGTGGCTGCGAGAGTTTAAAGACAG 1391
1329 CCCTGCGCGAGAGAGAACCCCGCGGTAAACAGAGAGCGAGAGCGCGCTCGCGCC 1388
1392 CAGCAGCTCCCTCACAGATTCCCAGAGTTTCTCTAAGTCAAGAGTCCAGCTGACAGAGAGCG 1451
1389 CACGAGAGAAAGAGAGAGAG---TGTTAAGATTAGAACCCAGCAGAGATTAACCTCAGCC 1445
1452 CAGTGAAGCAGTGTGTGAGAGGCCCGACACACATGGGCCAGGCCCTCGGGCCCCCGC 1511
1446 CCGGAGCCAAAGCAGAGCTCCCTGACTGGCAGTCCGCTCAGCTTTCGAGCTGTCTGTA 1505
1512 ACTCAGACTGTCCAGGCTTTTCCAGAGAGTCTCAGCTTCTCCCTCCAGCCGGA 1571
1506 CTGCGAGGGGGAGTGGGCCCTGCGCAACGTACCTGTCTCCCACTGAGCCCGGACAC 1565
1572 GAGCCCCCTGAGTGTGAGCCCTGACATTTCAAGTGTGCTT-----TGGGCCAGCTGA 1622
1566 CAGCAGGCTCTCTCCCTTGAATCAGCCGCTGTGAATGACAGGCTGCCCCCGGACGA 1625
1623 ACAGGCTCATCTCCCACTCTGAGGCCCTGTCTCTCAGTATGTGCGGAGCGGGATGGA 1682
1626 CGATGAGGATGAGGGCCGAGGAAAGATTCCTCACTGAATAAGGCTTCTCATGTACTTA 1685
1683 CTGCGAG---GAGCCACACACAGAGCGCTGCTGACAGAGGCTTCTCATATGTCAA 1739
1686 GGAAGTGTCTACACCGAGCGAATCTGAAGGATCTCGAAGTTATCATTCGTGTGTT 1745
1740 AGAGATTCCTGCTCAGAGACGAACATACCTCAAGGATTTAGAAAGTTATTAACGTTGTGTT 1799
1746 TCAGAGCAGATGAGACAAAGAGAGAGCCATGCGGAAACCTGAAABATCTCATATTTCC 1805
1800 CCGAGCCCAAGTGTGAAGAGAGAGAGCGCATGCTGCGACTGATGAGCGCTGCTTCTC 1859
1806 GAATTTTGAACCTTGGCAAAATTTCACTAATTTTCTCAAGAAATTTGCAACAGACT 1865
1860 CAACATCAATCCACTATATGATTTCCAGAGGCTTCTTGGCGAGGTGGAGCAGAGCT 1919
1866 TGCCCTGTGGAAAGGCGCTCAAAATGCCCCAATCGAG---ATTACAAAGATTCGCGGA 1922
1920 GGCACCTTGGGAAGGGCCCTCCAAAGGCCACACAAAGGAGTCAATCAACGAATCGGGGA 1979
1923 TGTCAATGCTGAAGAACATTCAGAGGCAATGAGACACTGGCGGCTCACTGTGAAACACAG 1982
1980 CATCTGTCTCAGAAACAGTGGCCAGTTAAAGAGTTTACAGCTATCTTCAAAAGACATGA 2039
1983 CGAGGCTTTGAGGCGCTGGAAGATGAAATCAAGAGCTCCCGCGGCTGAGAACTTCTG 2042
2040 CGAGGTCTTAACAGAACTGGAAGAGCTACCAAGCTGTGAAGAGTTGGAGGAGGTGA 2099
2043 CAGAGACTTTAGCTGCAAGAGGTGTATCACTTCAAGCTCAACACTTCTCTCGGCGC 2102
2100 CAAGGATTTGAGCTGAGAAAGGTGTGCTACTTGGCTCTCAACAGTGTCTGTGTAAGCC 2159
2103 ACTGACCGGCTCATATGCACTAACAGAGTCTTGAAGGGCTGTGCAAAACCAACCGCC 2162
2160 CATTCAGGGGCTGTGCACTACCGCTGTGTGCGCGCTATGCGGACATTTACAGCCC 2219
2163 GAGCCAGCCCACTTCAGGAGCTGCGAGCGCTTTTGGCAGAGATCAGAGATGTGTGC 2222
2220 CGGGACCATATGACTACGTGACTGCAATGAGCGCTTGAAGGCACTACAGAGGTACAC 2279
2223 ACAGCTCAACGATACGATGATCAAGATGAGAAATTTCCAGAAAGCTGACGAACCTCAAGA 2282

Db 2280 CACACTACAGACATTTCTATCCGGCTGGAGAACCTGCAAGCTAACGAGCTGACGCG 2339
Qy 2283 AGATTGATTGGCATTTGACATCTTGTGTTCCGGGAAGGAGTTCACTCGTGGGCGAG 2342
Db 2340 GAGCTGTGGGATAGAGAACTCATCTCTGGGAGGAGTTCACTCGTGAAGGCGCTG 2359
Qy 2343 CCAACAGAGCTCTCGGGGAGAGGGGCTCCAGAGCGCATTTCTCTCTGTTCAACAGCT 2402
Db 2400 CTTTCAACAGCTTCAAGAGAGGCTGACAGAGAGATTTTCTGTTCAAGATAT 2459
Qy 2403 CTTGCTATACAGAGCGGGGGCTGACGGCTCCATCACTTTAAAGTCCAGCGGAGCT 2462
Db 2460 GTTGCTGTACAGAGAGAGAGAGTTGACAGGACAGCACTTCGAGTCCGGGGCTCCT 2519
Qy 2463 CCGGCTCTATGAGCATGAGATTGAGAGAGAGAGAGAGAGTGGGGGCTGCGCTGCT 2522
Db 2520 TCCCTCTCAAGAGGATCTGTGTGAGAGAGATTAAGAGTGTCTGTTCAACAGCTGTT 2579
Qy 2523 GACCTCTCGGGGCGCAGCGGAGTCCATCATCTGTGGCGGCACTTCTCGGTCGAGATGGA 2582
Db 2580 CACACTTACGCGGCTCAAGAAACATCTGTGTGGAGCGCAGCATGCGCTGAGAGAAAG 2639
Qy 2583 GAAAGTGTGAGAGACATCCAGATGGCCATTTGACTGGCGAGAGAGAGAGAGAGCGCCG 2642
Db 2640 GAAAGTGTGAGAGACATCCAGATGGCCATTTGACTGGCGAGAGAGAGAGAGAGAGCG 2659
Qy 2643 CCGTGAATCTCTGGCCAGCA-----GCCCTCTGACACAGATCCCTGATGAGCGC 2696
Db 2700 CCGTGAATCTCTGGCCAGCACTGTGTGACATCTGCTCCCGCAGATCCCGCAGAGATAT 2759
Qy 2697 CGGCGGTGACAGAGATGAGAGATGACCTGAGCGGCTGCGGACATCTGCTGAGAGCGCA 2756
Db 2760 TCTG---GAGCAGAGATGAGAGATGATGCTCGGGGTGTCGAGCTCCTGAGAGCGCA 2816
Qy 2757 GCGCCCGCAGCGGGGAGACAAATGATGACATGTGTGCGACCGGACACAGCGCTCTC 2816
Db 2817 TGGCCAGACCGGGGCAACCAACATGACATGTGTGTGATGACGAGACACAGCGCTGTC 2876
Qy 2817 CATGTGATCTTACAGATGCGCACTGAGAAATCACTGTTGTTGAAACCTGCTGAGAAAT 2876
Db 2877 CAGGGGAGACCAAGTGTGAGCTGTGAGAAACAGCTTTCAGATATCTGCTAAGAAAT 2936
Qy 2877 CAAAAACACAAACGGGTGAGAGAGTGTGGGTGTTCAACAATCTGCTGCTGTTCTT 2936
Db 2937 CAAAAACATGATGAGTGTGAGAGAGTGTGGGTGTTCAACAATCTGTTGTTCTT 2996
Qy 2937 CTACAAATACACAGAGACATATCCCTTGGCAGCTGCTGTGCTGAGTACGCT 2996
Db 2997 CTACAAATACATCAAGATGATACCACTGCGCAGCTCCGCTGCTGAGCTACAGCGT 3056
Qy 2997 CACCATCTCTCTGAGTCCGAGAAATCCAGAAAGATCTGTTTCAAGTGTGACTTTCA 3056
Db 3057 GAGCATCTCCAGGGGAGCGGATGATACCAAGATCTGTTTCAAGTGTGACTTTCA 3116
Qy 3057 GTCCCAAGTCTACTTCTGAGGGGAGAGAGAGTCAAGTGTGAGAGAGTGAAGT 3116
Db 3117 ATCCCAAGTCTACTTCTGAGGGGAGAGAGAGTCAAGTGTGAGAGAGTGAAGT 3176
Qy 3117 GATCCGAGTGTGACAGAGCTGTGCTGCGAGCC 3151
Db 3177 GATCCAGGGGCGCAGAGCTCAACCGGAGGGGCC 3211

RESULT 7

US-10-802-432-16
; Sequence 16, Application US/10802432
; Publication No. US20040185489A1
; GENERAL INFORMATION:
; APPLICANT: Hoffmann-La Roche Inc.
; TITLE OF INVENTION: Transcriptional Activity Assay
; FILE REFERENCE: 21574
; CURRENT APPLICATION NUMBER: US/10/802,432

; CURRENT FILING DATE: 2004-03-17
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 16
; LENGTH: 3997
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: KIA0793
; LOCATION: (1)-(3997)
; OTHER INFORMATION: GenBank accession No. AB018336
US-10-802-432-16

Query Match 30.1%; Score 1037.4; DB 9; Length 3997;
Best Local Similarity 60.2%; Pred. No. 1,1e-305;
Matches 1864; Conservative 0; Mismatches 1186; Indels 45; Gaps 7;

Qy 69 GCGGACCCAGAGATCAAGACTGGGGGCCCGGAGAAATTCGGGAGTCACTTGGAGCG 128
Db 150 GCGAGTGCAGGGAGTGGCTTGGGTGCCAGACCCCTGTGGAGTTAGCACTTGAAGC 209
Qy 129 TGGACAGAAAGCGCCCGCCCAACACTTCAAGAAATCTGTGCTCATTAATCAAGTGT 188
Db 210 TGGGCACTCTCTTGGCCAGATGACAGAGAGAGCACTGCACTCAGATTAAGTGTCT 269
Qy 189 GAGTACACCCAGAGAGGAGCTTTGAAGTTCACAAAGAGTCCCGGAGAGTGTGCTGGA 248
Db 270 GAGCAACACCATGAAATATTTGACATGACCTTAATGCAATGGCAGATTAATGAGC 329
Qy 249 TGCAGTTTCAACCACTCTGAGAGAGTGAATTTTGGCTCGAGTTTCTGA 308
Db 330 ACAAGTGTGAAGCGTTTAACTGTGATAGATGATGATCTTGGGATGAGATTTCAAA 389
Qy 309 TCACAAAAAGATCAGGTGTGTGATCTCTAAACCATTTGAAACAGATTTGAAG 368
Db 390 TACTCAGTCTTCTGATTTGGCTTGAACCTAATGAACCATATTAAGCAATAGGAAG 449
Qy 369 GCCAAGCAAGTTTGTGTTAAATTTGTGTGAAATCTTTCGGCGGACACACACAACT 428
Db 450 GCCAAGAAATGTGTGTGCTTGGCTTGTGATTAATTTTTCACCTGATCTGTGTGCT 509
Qy 429 CCAAGAAAGATCACAAGGATCTGTTCCGCTGACAGTGAAGAGAGAGCTTGGCTCAAG 488
Db 510 ACAAGAAATATACAGATCTTGTGCTTGTGCACTTAAGAGAGAGCTGTGGAAGA 569
Qy 489 CAGTTGACGTGATGACACAGAGCAGCTCTTTGATTTCAACATTTGCAATCTGA 548
Db 570 GCGTTGACGTGTGCTGACACAGCGGCGCTTCTCAGTCCCATCTCTCAGTCGGA 629
Qy 549 GATGGGAGTTTGTGAGAGCTTGGACAGAGAGAGCTTAAGAGAAATTAATPACATAC 608
Db 630 AATAGAGATTAAGATGAAGAGCTGAGCCGAGAGAGAGCTCAAGATGAAGATTTGCG 689
Qy 609 TCAAGAGAGAGCTAG 668
Db 690 TGGCAGAGAGAGCTGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 749
Qy 669 ACCAGAGAGATGAGATTTCAAGTCTTGAAGATTTGCGGCTGAGAGATGATGAGAT 728
Db 750 ACTGTGAGTGTGATTTCAAGTGTGCTCAAGATTTGCTCAAGATTTGAGATGAGAT 809
Qy 729 CCGGTTGACCGGGGAG 788
Db 810 CAGATTTCAAGATGCTTCTGACAGAGAGAGAGAGAGAGATTAAGTGTGAGATTTCCAGAT 869
Qy 789 GGAATTTCTAGTGTTCAGAGGTTTCACTAAGATGATGCTTCAACTGGGCGCAAGTGG 848
Db 870 GGGTGTACTGTGTTCAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 929
Qy 849 GAAAGTGTGAG 908
Db 930 TAAAGTGTGAG 989

DB 3117 ATCCGACGTCTACTTCTCCGGGCTGAGCAAGTACATTTGAAAGGTGATGAGGT 3176
QY 3117 GATCCGACGTGACGACGAGCTCTGCTCGGAGCCC 3151
DB 3177 GATCCAGGGGGCCAGACGCTCAGCCCGGAGGGCCC 3211

RESULT 8

US-09-967-768A-104/C
Sequence 104, Application US/09967768A
Patent No. US2002015087A1
GENERAL INFORMATION:
APPLICANT: Augeneus, Meena
TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
FILE REFERENCE: 689290-72
CURRENT APPLICATION NUMBER: US/09/967, 768A
PRIOR FILING DATE: 2001-09-28
PRIOR APPLICATION NUMBER: US/60/236,109
PRIOR FILING DATE: 2000-09-28
PRIOR APPLICATION NUMBER: US/60/236,034
PRIOR FILING DATE: 2000-09-28
PRIOR APPLICATION NUMBER: US/60/236,111
PRIOR FILING DATE: 2000-09-28
NUMBER OF SEQ ID NOS: 325
SOFTWARE: PatentIn version 3.0
SEQ ID NO 104
LENGTH: 506
TYPE: DNA
ORGANISM: Homo sapiens
US-09-967-768A-104

Query Match 13.8%; Score 474.4; DB 3; Length 506;
Best Local Similarity 99.6%; Pred. No. 1.1e-133;
Matches 486; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 2824 GACTTCAGCATGCGACATGAGAAATCAGTTGTCTGAAACCTGCTGAGAAATTCAAAAAC 2883
DB 506 GACTTCAGCATGCGACATGAGAAATCAGTTGTCTGAAACCTGCTGAGAAATTCAAAAAC 447
QY 2884 AGCAACGGGTGCGAGAAAGCTGTGGGTGGTGTTCACAAACTTCTGCTGTTCTTACAAA 2943
DB 446 AGCAACGGGTGCGAGAAAGCTGTGGGTGGTGTTCACAAACTTCTGCTGTTCTTACAAA 387
QY 2944 TCACACACAGACAAATATATCCCTTGGCAGCTGCTCTGCTGGCTACTCTGCTACACATC 3003
DB 386 TCACACACAGACAAATATATCCCTTGGCAGCTGCTCTGCTGGCTACTCTGCTACACATC 327
QY 3004 CCTCTGAGTCCGAGAAATCCAGAAAGTACAGTGTTCAGAGCTGCACTTCAAGTCCAC 3063
DB 326 CCTCTGAGTCCGAGAAATCCAGAAAGTACAGTGTTCAGAGCTGCACTTCAAGTCCAC 267
QY 3064 GTCTACTTCTTCAAGGCGGAAAGCGAGTACAGTTCGAAAGGTGATGGAAGTATCCGC 3123
DB 266 GTCTACTTCTTCAAGGCGGAAAGCGAGTACAGTTCGAAAGGTGATGGAAGTATCCGC 207
QY 3124 AATGCCACACAGCTCTGCTCGGACACCCCACTGTTCAGCAACAAAGTCTCTGTGTAT 3183
DB 206 AATGCCACACAGCTCTGCTCGGACACCCCACTGTTCAGCAACAAAGTCTCTGTGTAT 147
QY 3184 TGATGGCCGAGACACACTGTTTCCGAGTGGCTGTTTCTGGAAGAGCTTTCCTTTCTT 3243
DB 146 TGATGGCCGAGACACACTGTTTCCGAGTGGCTGTTTCTGGAAGAGCTTTCCTTTCTT 87
QY 3244 CTGTATTAATGAAGCTGTGTAATTAACACCTGTCTGAAATCAAAAACATGGCTTCCC 3303
DB 86 CTGTATTAATGAAGCTGTGTAATTAACACCTGTCTGAAATCAAAAACATGG-TTCCC 28
QY 3304 AGCAGCTC 3311
DB 27 AGCAGCTC 20

RESULT 9

US-10-843-641A-6249/C
Sequence 6249, Application US/10843641A
Publication No. US2005006445A1
GENERAL INFORMATION:
APPLICANT: Avalon Pharmaceuticals, Inc.
TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
FILE REFERENCE: 689290-189
CURRENT APPLICATION NUMBER: US/10/843,641A
PRIOR FILING DATE: 2004-05-12
PRIOR APPLICATION NUMBER: US/09/873,367
PRIOR FILING DATE: 2001-06-05
PRIOR APPLICATION NUMBER: US/09/954,531
PRIOR FILING DATE: 2001-09-18
PRIOR APPLICATION NUMBER: US/09/954,456
PRIOR FILING DATE: 2001-09-25
PRIOR APPLICATION NUMBER: US/09/962,436
PRIOR FILING DATE: 2001-09-25
PRIOR APPLICATION NUMBER: US/09/962,832
PRIOR FILING DATE: 2001-09-25
PRIOR APPLICATION NUMBER: US/09/964,824
PRIOR FILING DATE: 2001-09-27
PRIOR APPLICATION NUMBER: US/09/967,768
PRIOR FILING DATE: 2001-09-28
PRIOR APPLICATION NUMBER: US/09/968,007
PRIOR FILING DATE: 2001-10-02
PRIOR APPLICATION NUMBER: US/09/969,347
PRIOR FILING DATE: 2001-10-02
PRIOR APPLICATION NUMBER: US/09/969,708
PRIOR FILING DATE: 2001-10-03
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 8447
SOFTWARE: PatentIn version 3.0
SEQ ID NO 6249
LENGTH: 506
TYPE: DNA
ORGANISM: Homo sapiens
US-10-843-641A-6249

Query Match 13.8%; Score 474.4; DB 10; Length 506;
Best Local Similarity 99.6%; Pred. No. 1.1e-133;
Matches 486; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 2824 GACTTCAGCATGCGACATGAGAAATCAGTTGTCTGAAACCTGCTGAGAAATTCAAAAAC 2883
DB 506 GACTTCAGCATGCGACATGAGAAATCAGTTGTCTGAAACCTGCTGAGAAATTCAAAAAC 447
QY 2884 AGCAACGGGTGCGAGAAAGCTGTGGGTGGTGTTCACAAACTTCTGCTGTTCTTACAAA 2943
DB 446 AGCAACGGGTGCGAGAAAGCTGTGGGTGGTGTTCACAAACTTCTGCTGTTCTTACAAA 387
QY 2944 TCACACACAGACAAATATATCCCTTGGCAGCTGCTCTGCTGGCTACTCTGCTACACATC 3003
DB 386 TCACACACAGACAAATATATCCCTTGGCAGCTGCTCTGCTGGCTACTCTGCTACACATC 327
QY 3004 CCTCTGAGTCCGAGAAATCCAGAAAGTACAGTGTTCAGAGCTGCACTTCAAGTCCAC 3063
DB 326 CCTCTGAGTCCGAGAAATCCAGAAAGTACAGTGTTCAGAGCTGCACTTCAAGTCCAC 267
QY 3064 GTCTACTTCTTCAAGGCGGAAAGCGAGTACAGTTCGAAAGGTGATGGAAGTATCCGC 3123
DB 266 GTCTACTTCTTCAAGGCGGAAAGCGAGTACAGTTCGAAAGGTGATGGAAGTATCCGC 207
QY 3124 AATGCCACACAGCTCTGCTCGGACACCCCACTGTTCAGCAACAAAGTCTCTGTGTAT 3183
DB 206 AATGCCACACAGCTCTGCTCGGACACCCCACTGTTCAGCAACAAAGTCTCTGTGTAT 147
QY 3184 TGATGGCCGAGACACACTGTTTCCGAGTGGCTGTTTCTGGAAGAGCTTTCCTTTCTT 3243
DB 146 TGATGGCCGAGACACACTGTTTCCGAGTGGCTGTTTCTGGAAGAGCTTTCCTTTCTT 87
QY 3244 CTGTATTAATGAAGCTGTGTAATTAACACCTGTCTGAAATCAAAAACATGGCTTCCC 3303

Db 86 CCGTATATGAAGCTGTGTAATAATTAACACCTGTGTGAAATCAAAAACAGG-TTCCC 28
Qy 3304 AGCAGCTC 3311
Db 27 AGCAGCTC 20

RESULT 10
US-09-764-868-51
; Sequence 51, Application US/09764868
; Patent No. US20020168711A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT232
; CURRENT APPLICATION NUMBER: US/09/764,868
; PRIOR APPLICATION DATE: 2001-01-17
; NUMBER OF SEQ ID NOS: 1510
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 51
; LENGTH: 1718
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-868-51

Query Match 11.7%; Score 402.4; DB 3; Length 1718;
Best Local Similarity 67.1%; Pred. No. 2.4e-111;
Matches 603; Conservative 0; Mismatches 286; Indels 9; Gaps 2;

Qy 2260 CAGAAGCTGACGAACTCAAGAAAGATTGATGGCATTGACAAATCTGTGTCCGGGA 2319
Db 13 CAGAACTAACGAGAGCTGACGCGGACCTGTGTGGCATAGAACTCATTTGCTCTGAC 72

Qy 2320 AAGGAGTTTCCTGCTGCGCAGCCTCAGCAAGCTTCGCGGAAAGGAGCTCCAGACGCG 2379
Db 73 AAGGAGTTTCCTGCTGCGCAGCCTCAGCAAGCTTCAGCAAGAGGAGCTCCAGACGAGG 132

Qy 2380 ATGTTCTTCTGTTTCAACAGAGTCTGCTATACAGAGCCGGGGGCTGACGGCTTCAT 2439
Db 133 ATGTTCTTCTGTTTCAACAGAGTCTGCTATACAGAGAGGAGTTGACAGGAGC 192

Qy 2440 CAGTTTAAAGTCACAGGAGAGCTCCGCTCTATGGCATGACGATTGAGAGAGCAAGAC 2499
Db 193 CACTTCCGATCCGGGGGCTCTTCCCTCCAAAGGATGCTGTGAAAGAAATGATTAAC 252

Qy 2500 GAGTGGGGGGTGCCTGACCTGCTCCGCGGAGCAGGAGCATCATCTGTGCGC 2559
Db 253 GAGTGGTCTGTTCCACACTGTTTCAACATCTACGCGGCTCAGAAAACATCGTGTGCGC 312

Qy 2560 GCCAGTTCTCGGTCCGAGATGAGAGAGTGGATTGAGGAATCCAGATGGCCATTGACCTG 2619
Db 313 GCCAGATCTCGGTGAGAGAGAGTGGATTGAGGAATCTGAACTCCGCGATCCAGAGA 372

Qy 2620 GCGAGAAAG 2673
Db 373 GCGAG 432

Qy 2674 AACAAAGTCCCTGAGTGAAGCCACCGCGCTGACAGAGATCAGAGATGACCTGAGCGCC 2733
Db 433 CCGAGATCCCCCAACAGAGTATCTCTG---GAGCAGAGATCAGAGATGATCTCGGGGT 489

Qy 2734 TGGCGACATCGCTGAGAGGCGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2793
Db 490 GTCGCGAGCTCCCTGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 549

Qy 2794 TGGCAGCCGACAG 2853
Db 550 TGGTACCGAG 609

Qy 2854 TCTGGAACCTGTGTGAGAAATTCAAAACAGCAACGAGTGGAGAGAGTGTGTGTGTG 2913

Db 610 TCAGATATCTGTAAAGAAAGTTCAAAAACAGTACGTGCTGGAGAAAGCTGTGCTGTC 669
Qy 2914 TTCAAACTTCTGCTGTTTCTTCTTCAAAATACACCGAGACATCATCCCTTGGCAGC 2973
Db 670 TTACCAACTTCTGTTTCTTCTTCAAAATACATCAGATGATACCTCCAGTCCAGC 729

Qy 2974 CTGCTGTGCTGAGTACCTGCTACACATCCCTCTGTGATCCGAGAAATCCAGAAAGC 3033
Db 730 CTCCGCTGCTGAGGCTTACAGGTGAGCATCCCGAGGAGGCGGATGAGATACAAAGAC 789

Qy 3034 TACGTGTTCAAGCTGACCTTCAAGTCCACAGTCTTACTTCTTACGAGGCGAAAGCAGTAC 3093
Db 790 TATGTTTCAAGCTTCAAGTCTTCAAAATCCACAGTCTTACTTCTCCGGCTGAGAGCAAGTAC 849

Qy 3094 ACGTTCAAGAGTGGATGAGAGATGATCCGAGTGCACAGCTGCTGCTCGGAGCC 3151
Db 850 ACGTTCAAGAGTGGATGAGAGATGATCCAGGAGGCGCAGAGCTCAGCGGAGGCGCC 907

RESULT 11
US-09-918-995-8794
; Sequence 8794, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8794
; LENGTH: 393
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(393)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-8794

Query Match 11.2%; Score 385.6; DB 3; Length 393;
Best Local Similarity 99.0%; Pred. No. 1.5e-106;
Matches 388; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 910 TACAGATTAACCTTGAATTCCTGATGCGCAGTGGAGATTTCTGCAAGCTCTTGTGAAA 969
Db 1 TACAGATTAACCTTGAATTCCTGATGCGCAGTGGAGATTTCTGCAAGCTCTTGTGAAA 60

Qy 970 ATCTGTGTTGAACATCATGCTCTTCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1029
Db 61 ATCTGTGTTGAACAGCATGCTCTTCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120

Qy 1030 CCGGCTCTTGAAGCGGAGGATCATCATTTGAGTTCAAGTGTGAGTCAAGAGAGAGTT 1089
Db 121 CCGGCTCTTGAAGCGGAGGATCATCATTTGAGTTCAAGTGTGAGTCAAGAGAGAGTT 180

Qy 1090 CTGCACTATGTTAAAGAGAGAGATTAAGAGTGCAGTTTGAAGAGAGAGAGAGAGAG 1149
Db 181 CTGCACTATGTTAAAGAGAGAGATTAAGAGTGCAGTTTGAAGAGAGAGAGAGAGAG 240

Qy 1150 ATTCACTTATCCGAGAGCTTGTCTTCAAGCTTCAAGAGAGAGAGAGAGAGAGAGAGAG 1209
Db 241 ATTCACTTATCCGAGAGCTTGTCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300

Qy 1210 CAGTCTGAG 1269
Db 301 CAGTCTGAG 360

Qy 1270 AGCTGCCGCGAG 1301

Search completed: May 22, 2006, 11:24:23
Job time : 3957.83 secs

```
QY 1755 AGTGAACAAAGAGAGCGCATCCGGAACACTGAAAAGTCTCATATTTCCGAATTTTGA 1814
DB 129 AGTGAGCAAAAGAGAGTCCATGCCGAAACCTTGAGAAATCTCATATTTCCGAATTTTGA 188
QY 1815 ACCTTTGACAAATTTTCATATTTTCTCAAGGAAATTTGAGCAAGCATTTGCCCTGTG 1874
DB 189 ACCTTTGACAAATTTTCACACATATTTTCTCAAGGACATTTGAGCAACGACTTGCCTGTG 248
QY 1875 GGAAGGCGCTGCAATGCGCAATCA--GAGATTAACCAAGATCGCGATGTCATGCT 1931
DB 249 GGAAGGCGCTGCAATGCGCAATCAAGAGATTAACCAAGATCGCGATGTCATGCT 308
QY 1932 GAAGAACATTTGAGGCGATGAAGCACTTG-GCGGCTCACTGTGAAAGCAGAGGCT 1990
DB 309 GAAGAACATTTGAGGCGATGAAGCACTTGAGCTGCTCATTTGTGAAAGCAGACTAGGCC 368
QY 1991 TGGAGGCGCTGGAAGATGAATCAAGAGCTCC 2022
DB 369 TGGAGGCGCTGGAAGATGAATCAAGAGCTCC 400
```

```
RESULT 15
US-10-062-674-58
; Sequence 58, Application US/10062674
; Publication No. US20040005559A1
; GENERAL INFORMATION:
; APPLICANT: Loring, Jeanne F.; Kaeber, Matthew R.
; TITLE OF INVENTION: MARKERS OF NEURONAL DIFFERENTIATION AND MORPHOGENESIS
; FILE REFERENCE: PA-0026-1 CIP
; CURRENT APPLICATION NUMBER: US/10/062,674
; CURRENT FILING DATE: 2002-01-30
; PRIOR APPLICATION NUMBER: US 09/525,102
; PRIOR FILING DATE: 2000-07-24
; NUMBER OF SEQ ID NOS: 2217
; SOFTWARE: PERL Program
; SEQ ID NO 58
; LENGTH: 437
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: GenBank ID No. US20040005559A1 g2305969
US-10-062-674-58
```

Query Match 8.3%; Score 286.8; DB 7; Length 437;
Best Local Similarity 88.1%; Pred. No. 3e-76;
Matches 312; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

```
QY 2848 CAGTTGTCTGGAACCTGCTGAGAAATTCAAAACAGCAACGCGGTGCGAAGAGCTGTG 2907
DB 1 CAGCTCTCTGGGAACCTGCTGAGAAATTCAAAACAGCAACGCGGTGCGAAGAGCTGTG 60
QY 2908 GTGAGTTCACAACCTTCTGCTGTTCTTCAAAATCACACAGAGCAATCATCCCTT 2967
DB 61 GTGATTCACCAACTTCTGCTGTTCTTCAAAATCACACAGAGCAATCATCCCTT 120
QY 2968 GCCAGCTGCTCTGCTGCTGCTACTGCTCACCATCCCTCTGAGTCGAGAAATCCAG 3027
DB 121 GCCAGCTGCTCTGCTGCTGCTACTGCTCACCATCCCTCTGAGTCGAGAAATCCAC 180
QY 3028 AAAAGCTAGTGTCAAGCTGCACTTCAAGTCCCAAGTCTACTACTTCAAGGCGGAAAG 3087
DB 181 AAGGACTATGTGTCAAGCTGCACTTCAAGTCCCAAGTCTACTACTTCAAGGCTGAAAGC 240
QY 3088 GAGTACAGCTTGAAAGGTGATGAGTATCCGAGTCCAGCAAGCTCTGCTGCGA 3147
DB 241 GATATCACTTTGAAAGGTGATGAGTATCCGAGTCCAGCAAGCTCTGCTGCG 300
QY 3148 CCCACGCTGTGAGCCACAAAGTCTCTTGTATTTGATGCGCGACACTC 3201
DB 301 GCCCAGCTTGTAGTACAAAGATCTCACTGTACTGAGGCTGACAGCGCC 354
```



```

QY 345 ACCCATTTGAAAACAGATTAGAAAGCCAAAGCAGTTGTTGTTAAAGTTTGTTGTAAGT 404
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 978 ABAATTAATAAAAGGAGGTTGCGG---TGTCCTTGGAATTTTACCTTAATGTAAGTT 1034
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 405 CTTTCGGCTGACACACACAACTCCAGAGAACTACAGAGTACCTTGCGCTGCA 464
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1035 TTATCCACCTGACCCGACAGTAAACAGAAACATACAAAGATATTTATGCTTCA 1094
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 465 GGTGAAGCAGGACCTGGCTCAAGGAGGTTGACGTTAATGACACGAGGCGCTCTT 524
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1095 GCTTGGGAGAACATGTTGACAGAGCTTCCCTGCTTCTTGCACTTACATTAAT 1154
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 525 GATTTACACATTTGCAATCTGAGATGGGATTTTGAAGACCTTGAACAGAGCA 584
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1155 AGTTCTTACCATCATCAGCTGGAAGTGGAGACTGACCCAGAACTCATGGGCTGA 1214
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 585 CTTAGCAAAAATTAATACATACC-----TCAGCAAGACGCACTGAGAGCAAAATCGT 638
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1215 TTATGTTAGTGAATTTTAACTGGCCCCGAATCAGACCAAGAACTTGAAGAGAGTCA 1274
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 639 GGAATTTCCCAATTAACCAATTTGAGCAAAACACAGAGATCAGATTTCCAGCTCTAGA 698
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1275 GGAATGCTAATAGTACATACAGGTCATGACTCCAGCTCAGCTGACTTGAAGTTCTGA 1334
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 699 GATTGCCCCGCTGAGATGATGAAATCCGTTGCAACCCGCGCAAGAGAGAGAG 758
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1335 GAAATGCAAAAAGTTGTATGATGAGATGATGATCTTCAAAAGCAAGAGACTGGAGAG 1394
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 759 CACGAAGATCAATCTGGCCGCTTGCCAAACAGGAAATTCAGTGTTCAGAGTTCACTAA 818
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1395 AGTAGATATCATCTTAGGTCTGCTCTAGTGCCCTTGCTGTTTCAAAAGATTAAGCTGAG 1454
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 819 GATCAATGCTTCAACTGAGGCCAAGGTGGGAGAGCTGAGCTTCAAGAGAGAGCGCTTCT 878
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1455 AATTAAACCGCTTCCCTGGGCCAAAGTGTGAAGATTTCTTAATAACGTAAGCTTTT 1514
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 879 CATCAAGCTCCGCGCAGATGCCAATAGTCCGTAACAGATACCTTGAATTCCTGATGCGC 938
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1515 CATCAAGATTCGGCCCTGGAGAGCAAGAGCATATGAAGTACATGAGATCAAACTTCC 1574
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 939 CAGTGGGATTTCTGAAATCTTCTGGAATAATCTGTGTAACATCAATGCTTCTTTG 998
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1575 CAGTTAACGAGACGCTAAGAAATTAAGAAAGCTGTGTGAACATCACACGTTTTTCAG 1634
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 999 ACTTTTGAAGACCCAAACCAAGCCCAAGCCGCTCTTTAAGCCGGGGATCATTT 1058
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1635 A---TTGACATCTACAGACACCATTTCCAAAAGCAAAATTTCTTGCGCTAGATCCAAAT 1691
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1059 TCGGTTCACTGGTCCGACTCAGAAAGCAG 1086
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1692 TCGATTCAGTGGCCGGAATCAAGCTCAG 1719
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 2
US-10-473-173-56
; Sequence 56, Application US/10473173
; Publication No. US20060088823A1
; GENERAL INFORMATION:
; APPLICANT: VAN ANDEL INSTITUTE
; TITLE OF INVENTION: Microarray Gene Expression Profiling in Clear Cell Renal Cell
; FILE REFERENCE: 38345-170094
; CURRENT APPLICATION NUMBER: US/10/473,173
; CURRENT FILING DATE: 2003-09-29
; PRIOR APPLICATION NUMBER: US 60/279,411
; PRIOR FILING DATE: 2001-03-29
; NUMBER OF SEQ ID NOS: 498
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 56
; LENGTH: 3953
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-473-173-56

```

```

Query Match 1.7%; Score 56.8; DB 6; Length 3953;
Best Local Similarity 46.6%; Pred. No. 5.5e-06;
Matches 218; Conservative 0; Mismatches 247; Indels 3; Gaps 1;

QY 178 ATCCAGATCTGAGATGACACCCAGAGGACATTTGAAGTTCCAAAGAGCTCTGGAAG 237
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 284 ATCCAACTGCTTAATAAGAGTTGTGAGTTTCAACCTGCTCGTGAAGACACTGGCCAG 343
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 228 GTGCTGCTGAGATGACGTTTGCAACCACTCAACCTCGTGAAGAGTGAATTTTGGCCTC 297
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 344 GAAAGCTTGAGGCGGTGCGCCAGAGGCTGAGGCTGCGAGGTCACTTACTTCAAGCTTC 403
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 298 GAGTTTCTGATCAAAAAGATCAAGGTGCTGAGTCTGATCTCTTAAACCCATTTGAAA 357
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 404 TGCTACTACACACAAATATCAGCGCGGTGGTGAATTTGAAAAACCTTTGAAGAAAG 463
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 368 CAGAT---TAGAAGGCCAAAGCAGTTGTTGAAGTTGTGAGAAATTTCTTCCGCT 414
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 464 CAGCTGATTAATATGCAATTTGAACTACCGCTATTTTGGAGTGTGTTTATGTGCT 523
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 415 GACCAACACAACTCCAAAGAACTCAAGATGATCAGTCTGCTGCGCTGAGGTGAAGAG 474
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 524 TCAATTTCTCAGCTGACAGAGATTAACAGATATCAAGTATTAATGCACTGAAGAAA 583
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 475 GACTTGCTCAAGCAGGTTGACGTGTATGACACACAGCAGCTCTCTTGAATTTACAC 534
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 584 GATATCTTGAAGAGAAATTCCTTGTACCTTAAGAAAGAAATTCAGCTACAGAGCTTA 643
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 555 ATTGTCAATCTGAGATTTGGGATTTTATGAAGCTTTGAGACAGAGACACTTACAAA 594
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 644 GCTGTTCAAGGAGATTTTGTGACTTGTATCATATGAATCCAGAGACTTTCTTCAGAAA 703
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 595 AATAAATCATACCTCAGCAAGAGCAGCTAGAGACAAATGCTGAA 642
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 704 TTTCCTTGTCTTCTTGAGATGCTTACAAATGAAGAAATGTAAGAA 751
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 3
US-10-488-619-2736
; Sequence 2736, Application US/10488619
; Publication No. US20060099578A1
; GENERAL INFORMATION:
; APPLICANT: Greenlee, Winner and Sullivan, P.C.
; TITLE OF INVENTION: Human Mitochondrial DNA Polymorphisms, Haplogroups, Associations &
; FILE REFERENCE: 98-01 WO
; CURRENT APPLICATION NUMBER: US/10/488,619
; CURRENT FILING DATE: 2004-03-01
; NUMBER OF SEQ ID NOS: 3040
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2736
; LENGTH: 646
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-488-619-2736

```

```

Query Match 1.2%; Score 41.8; DB 6; Length 646;
Best Local Similarity 49.8%; Pred. No. 0.025;
Matches 106; Conservative 0; Mismatches 107; Indels 0; Gaps 0;

QY 2055 GCTGCAAGAGTGTGTTACTTACCGCTCAAGACCTTCTCTGCGGCACTGACCGGCT 2114
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 73 GCTGCAAGAGATGATGACATCTGCTGATGGGCTTTCTGCTGACCCCGGTGCAAGAT 132
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2115 CATGACATACAGACAGGTCCTGAGAGGCTGTGCAAAACACACCCCGGAGCAGCCGA 2174
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 133 CTGCAAGTATCCCTCCAGCTGCGAGAGCTGTCAAGTACACACCCCGGAGCAGGGA 192
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2175 CTTGAGGACTGCGGAGCGCTTTGGAGAGATCAAGAGATGTGGACACAGCTCCAGCG 2234
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 193 CTTTAAGATGTGAAGCTGCTTGCAGCGCATGAAGATGTGGCCAGCTCATACGA 252
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```


Qy		Db
2235	TACGATGATCAAGATGGAGAATTTCAGAGCT	2267
253	ACGGAAACGAGACTTGAATAACATCGACAAGAT	285

RESULT 4
US-10-473-173-32

Sequence 32 Application US/10473173
Publication No. US20060088823A1
GENERAL INFORMATION:
APPLICANT: VAN ANDEL INSTITUTE
TITLE OF INVENTION: Microarray Gene Expression Profiling in Clear Cell Renal Cell
TITLE OF INVENTION: Carcinoma: Prognosis and Drug target Identification
FILE REFERENCE: 38345-170094
CURRENT APPLICATION NUMBER: US/10/473,173
CURRENT FILING DATE: 2003-09-29
PRIOR APPLICATION NUMBER: US 60/279,411
PRIOR FILING DATE: 2001-03-29
NUMBER OF SEQ ID NOS: 498
SOFTWARE: Patentin version 3.2
SEQ ID NO 32
LENGTH: 37426
TYPE: DNA
ORGANISM: Homo sapiens
US-10-473-173-32

Query Match	1.2%	Score	41.4	DB	6	Length	37426
Best Local Similarly	54.2%	Pred. No.	0.36				
Matches	84	Conservative	0	Mismatches	71	Indels	0
						Gaps	0

QY	253	11233	QY	313	11293	QY	373	11353
GTGTGCAACACCTCAACCTCGGGAAGGAGCATATTTGGCCCTCGAGTTTCCATATCAC	GTGTGCAACACCTCAACCTCGGGAAGGAGCATATTTGGCCCTCGAGTTTCCATATCAC	GTGTGCAACACCTCAACCTCGGGAAGGAGCATATTTGGCCCTCGAGTTTCCATATCAC	AAAAAAGTACGGGTGTGGCTGAGATCTCTTAAACCATATGTGAAACAGATTAGAAAGGCCA	AAAAAAGTACGGGTGTGGCTGAGATCTCTTAAACCATATGTGAAACAGATTAGAAAGGCCA	AAAGGTAAAGTTATATGGCTTAAACCTGAGAAACCGGATCTCTCCAGCAGATGATGGGCTTA	AAGCACGTTGTGTAAAGTTGTGTGGAATTCCT	AAGCACGTTGTGTAAAGTTGTGTGGAATTCCT	GCCTCTTACAGGCTTAAACCTTGAAGTCACAGTTCCT
11387	11387	11387	372	372	11352	407	407	11387

RESULT 5

```

US-11-325-764-1
/ Sequence 1, Application US/1135764
/ Publication No. US20060099640A1
/ GENERAL INFORMATION:
/ APPLICANT: Jegla, Timothy James
/ APPLICANT: Witzel, Julie Dickson
/ APPLICANT: ICAGEN, Inc.
/ TITLE OF INVENTION: S102 and S104, Novel Potassium Channel Proteins from
/ TITLE OF INVENTION: Human Brain
/ FILE REFERENCE: 018512-006810US
/ CURRENT APPLICATION NUMBER: US/11/325,764
/ CURRENT FILING DATE: 2006-01-04
/ PRIOR APPLICATION NUMBER: US/09/921,159
/ PRIOR FILING DATE: 2001-08-01
/ PRIOR APPLICATION NUMBER: US 60/249,112
/ PRIOR FILING DATE: 2000-11-15
/ NUMBER OF SEQ ID NOS: 38
/ SOFTWARE: Patentin Ver. 2.1
/ SEQ ID NO 1
/ LENGTH: 3708
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (1)..(3708)
/ OTHER INFORMATION: human S102 potassium channel alpha subunit
/ US-11-325-764-1

```

Query Match 1.28; Score 40.4; DB 7; Length 3708;

Best Local Similarity 58.2%; Pred. No. 0.17;
Matches 71; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY	Db	QY	Db
2099	1917	2154	2036
CTGGGGGCACATGCAACGGGCTCATGCACTAACAACAGAGTCTGTGGAGCGGCTGTGAAACA	CCGCGTCCCGCTGACACACATCATCGCTCCATATGATGACATGACCTCTGACGGGACACAGA	CCACCTCCGCGGACGCCACGCGCACTTTCAGGGAATGTCGCGAGCCGCGCTTGTGCAAGATCACGGA	GCACCTCCGCTTACGACAGAGCCGCGCGTGGGGGCGGGGGGACGACAAGCTGTGGCACTGCGCCACGGA
1978	1978	2213	2036

QY	2214	GA	2215
Db	2037	GA	2038

RESULT 6
US-10-505-928-744

```

? Sequence 744, Application US/10505928
? Publication No. US2006008653A1
?
? GENERAL INFORMATION:
?
? APPLICANT: Ludwig Institute for Cancer Research et al
?
? TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
? FILE REFERENCE: 28967/39178
? CURRENT APPLICATION NUMBER: US/10/505,928
? CURRENT FILING DATE: 2004-08-27
? PRIOR APPLICATION NUMBER: US 67/363,019
? PRIOR FILING DATE: 2002-03-07
? NUMBER OF SEQ ID NOS: 866
? SOFTWARE: PatentIn 3.2
?
? SEQ ID NO 744
?
? LENGTH: 2367
?
? TYPE: DNA
?
? ORGANISM: Homo sapiens
?
? US-10-505-928-744

```

Query Match	1.1%	Score	38.4	DB	6	Length	2367
Best Local Similarity	46.3%	Pred. No.	0.47				
Matches	126	Conservative	0	Mismatches	145	Indels	0
						Gaps	0

Oy	1915	ATCGGCGCATGTGCATGCTGGAAAGACATTCAAGGGCACATCGTGGGGCTCACCCTGTGG	1974
Db	448	ATTCCTCGAGCTCCGCAAGATGAATTTAAGAAGGATCTCCAGCACCTCTTAGCGCTCTGTCTGT	507
Oy	1975	AAGCACACGAGGCGCTTGAGAGGCCCTCGAGAAATCAAGACTCCCGCGGCTGGAG	2034
Db	508	GTAACAACAAGATCTCCAGATTCATGAGAAAGGCTTCAGCCCACTGGGAAGCTGGAC	567
Oy	2035	AACCTTCTGCAGAGACTTTGAGCTGCAGAGGTGTGTTACTTAACCGCTTCACAACCTTCTC	2094
Db	568	AAGCTCTACATCTTCCAAGAACCACTCGTGGAGATCCCGCCCAACTTAACCGACTCTCCTG	627
Oy	2095	CTGCGGCGCATCGACCGGCTCATGCATAACAAGAGGCTCTGAGACGGCTGTGCAACAC	2154
Db	628	GTGAGAGCTCCGCATTCACAGCAAACCGCATCCGCAAGGTGCCAAGGGAGTTCACAGCGGG	687
Oy	2155	CACCCGCCGAGCCAGCGCGCACTTCAGGAGCTG	2186
Db	688	CTCCGGAACATGAATCGATCGAAGATGGCGG	719

RESULT 7
US-11-14

Sequence 189, Application US/11145307A
Publication No. US20060094035v1
GENERAL INFORMATION:
APPLICANT: Arcurus Bioscience, Inc.
APPLICANT: Exlender, Mark G.
APPLICANT: Ma, Xiao-Dun
TITLE OF INVENTION: Identification of Tumors
FILE REFERENCE: 022041-002020/US
CURRENT APPLICATION NUMBER: 2005/11/145, 307A
CURRENT FILING DATE: 2005-06-03

/ PRIOR APPLICATION NUMBER: US 60/577,084
/ PRIOR FILING DATE: 2004-06-04
/ NUMBER OF SEQ ID NOS: 268
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 189
/ LENGTH: 2711
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-11-145-307A-189

Query Match 1.1%; Score 36.4; DB 7; Length 2711;
Best Local Similarity 56.8%; Pred. No. 1.8;
Matches 67; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 195 CACCCAGAGGAGATTGAGATTCACAAAGAGCTCTGGGAAGGTGCTGTGATGACGT 254
DB 500 CAGCCGAGGAGCGGTCTCGATCTCCAGCAAACTGTGAGGTTCTGGCGTTTGTAGT 441
QY 255 TTGCAACCACTCAACCTGTGTGAGAGTATTTTGGCTTCTGATTCAC 312
DB 440 CCCCAATCTTCTCCACCACTTGGGGGTGGCCACTTGGGCTTGGAGCCCTATCAC 383

RESULT 8
US-10-505-928-380/c
/ Sequence 380, Application US/10505928
/ Publication No. US20060088532A1
/ GENERAL INFORMATION:
/ APPLICANT: Ludwig Institute for Cancer Research et al.
/ TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
/ FILE REFERENCE: 28967/39178
/ CURRENT APPLICATION NUMBER: US/10/505,928
/ PRIOR FILING DATE: 2004-08-27
/ PRIOR APPLICATION NUMBER: US 60/363,019
/ PRIOR FILING DATE: 2002-03-07
/ NUMBER OF SEQ ID NOS: 866
/ SOFTWARE: PatentIn 3.2
/ SEQ ID NO 380
/ LENGTH: 1323
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-505-928-380

Query Match 1.0%; Score 35.6; DB 6; Length 1323;
Best Local Similarity 42.8%; Pred. No. 2;
Matches 179; Conservative 0; Mismatches 239; Indels 0; Gaps 0;

QY 2322 GGAATTCATCCGCTCTGGGCAAGCTCTCGGGAAGGGCTCCAGACGGCAT 2381
DB 957 GGGGCTCATGGCAGCGGGTTTCGACGACTGTCTGGCAATGAAGACCAAGGGGCGAG 898
QY 2382 GTTCTTCCGTTCAACGAGTCTCTATACACAGCGGGGGCTGACGGCTCCCAATCA 2441
DB 897 CCAACACAGCTGGCCACCAACATGATCCCAAGAACTAAGCCATATCTCCACCTCGGA 838
QY 2442 GTTTAAAGTCAAGGGCAAGCTCCGCTCTATGSCATGACGATTTAGAGAGAGCA 2501
DB 837 GTCCCGGGGACGCTGTGGGGCGGCTCTGCGCGTGTAGACGTGCAAGGGTGGCCAC 778
QY 2502 GTGGGGGGTCCCGCACTGCTGACCTCCGGGGCCAGCGGAGTCCATCATGTGGCGGC 2561
DB 777 GCTGACCGGTTCAGAGGAAAGGACAGCCGACCGAGAGGGCGCCACGATGAGAAAG 718
QY 2562 CAGTTCTCGTCCGAGATGGAAGAGTGGGTGAGGACATCCAGATGGCCATGACCTGGC 2621
DB 717 CAGCCGAAAGGCGACGTCCCGGACTCGCGCCGACGCTCAGGAACACACAGGACCCCG 658
QY 2622 GGAGAAAGACAGAGCCCGCTGATGTTCTGTGSCACAGAGCCCTGACAAACATGTC 2681
DB 657 GATTTGCAAGGTGTAGCAACCAAGCCGCAAGGGGCAAGGGCCCAAGGCGGCGG 598
QY 2682 CCCTGATGAAGCCAGCGGAGCTGACAGAGATGAGAGTACGTAGGCGCTCGCGC 2739

DB 597 CGCCCAACCAAGCCCAAGGTGGCCCAAGCGCGGCTGCGAGGCGAGCCCGGCGGC 540

RESULT 9
US-10-511-455-1
/ Sequence 1, Application US/10511455
/ Publication No. US20060088835A1
/ GENERAL INFORMATION:
/ APPLICANT: Blackard, Benjamin Simon
/ APPLICANT: Blackwood, Douglas
/ APPLICANT: Porteous, David
/ APPLICANT: Muir, Walter John
/ APPLICANT: Muir, Walter John
/ APPLICANT: More, Ole
/ APPLICANT: Ewald, Henrik Lykke
/ TITLE OF INVENTION: SCHIZOPHRENIA ASSOCIATED GENES
/ FILE REFERENCE: 9013.63
/ CURRENT APPLICATION NUMBER: US/10/511,455
/ PRIOR FILING DATE: 2004-10-05
/ PRIOR APPLICATION NUMBER: PCT/GB03/001543
/ PRIOR FILING DATE: 2003-04-07
/ PRIOR APPLICATION NUMBER: GB0207902.8
/ PRIOR FILING DATE: 2002-04-05
/ PRIOR APPLICATION NUMBER: GB0207904.4
/ PRIOR FILING DATE: 2002-04-05
/ PRIOR APPLICATION NUMBER: GB0207900.2
/ PRIOR FILING DATE: 2002-04-05
/ PRIOR APPLICATION NUMBER: GB0207901.0
/ PRIOR FILING DATE: 2002-04-05
/ PRIOR APPLICATION NUMBER: GB0227734.1
/ PRIOR FILING DATE: 2002-11-28
/ NUMBER OF SEQ ID NOS: 94
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 1
/ LENGTH: 4773
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-511-455-1

Query Match 1.0%; Score 35.4; DB 6; Length 4773;
Best Local Similarity 44.5%; Pred. No. 4.8;
Matches 141; Conservative 0; Mismatches 176; Indels 0; Gaps 0;

QY 2459 AGCTCCCGCTTANGCATGACGATTTAGAGAGAGCAAGAGTGGGGGTGCCCACT 2518
DB 2219 AACTCCAGCTGATAGAGGCTGTGATGATGACAGAAAGCACTTCTGATGACCTGC 2278
QY 2519 GCTTACCTCCGGGGCCAGCGGCACTCATCATGTGCGCCCACTTCTCGGTCCAGA 2578
DB 2279 ACATGACATGTGTGAGGAGGACGACCAACAGGCCATGCAATTCACAGCTAGCTGTC 2338
QY 2579 TGAAGAAGTGGGTGAGGACATCCAGATGGCCATTGACCTTGGCGGAGAAAGACAGAGCC 2638
DB 2339 AGCAGAAAGACGACGAGAGCGGTGGGACCAAGATACAGCCATCTTGTCCAAAC 2398
QY 2639 CGCCCTGAGTTCTTGGCCAGACAGCCCTCTGACAAAGTCCCTGATGAAGCCACCG 2698
DB 2389 AGCAGAAAGACGAGCGGTGGGCGGACGAGAGGACCCGTATGACGAGAGCT 2458
QY 2699 CGGCTGACCAAGAGTCAAGATGACCTGAGGCGCTTCCGCGACATCGCTGAGAGCGCAAG 2758
DB 2459 CGGAGCAAGAAACATGCGGACGACGACCGCATCTCCAAACCGCTGGCGGGGACGA 2518
QY 2759 CCCGCAACGCGGCAAC 2775
DB 2519 GGAAGCTCAGCTGCAAC 2535

RESULT 10
US-10-473-173-50
/ Sequence 50, Application US/10473173
/ Publication No. US20060088823A1
/ GENERAL INFORMATION:
/ APPLICANT: VAN ANDEL INSTITUTE

```

GENERAL INFORMATION:
APPLICANT: Millennium Pharmaceuticals, Inc.
APPLICANT: Rosenfeld, Julie Beth
APPLICANT: S110s-Santiago, Immaculada
TITLE OF INVENTION: METHODS AND COMPOSITIONS IN TREATING
TITLE OF INVENTION: PAIN AND PAINFUL DISORDERS USING 9949, 14230, 760, 62553,
TITLE OF INVENTION: 12216, 17719, 41897, 47174, 33408, 10002, 16209, 314, 636,
TITLE OF INVENTION: 27410, 33260, 619, 15985, 69112, 2158, 224, 615, 44373,
TITLE OF INVENTION: 95431, 22245, 2387, 16658, 55054, 16314, 1613, 1675, 9569 OR
FILE REFERENCE: MPI02-027P1R0NANIM
CURRENT FILING DATE: 2005-12-20
CURRENT FILING DATE: 2005-12-20
PRIOR APPLICATION NUMBER: US/10/369, 022
PRIOR FILING DATE: 2003-02-19
PRIOR APPLICATION NUMBER: US 60/360, 495
PRIOR FILING DATE: 2002-02-28
PRIOR APPLICATION NUMBER: US 60/370, 121
PRIOR FILING DATE: 2002-04-04
PRIOR APPLICATION NUMBER: US 60/373, 010
PRIOR FILING DATE: 2002-04-16
PRIOR APPLICATION NUMBER: US 60/373, 908
PRIOR FILING DATE: 2002-04-19
PRIOR APPLICATION NUMBER: US 60/377, 717
PRIOR FILING DATE: 2002-05-03
PRIOR APPLICATION NUMBER: US 60/379, 949
PRIOR FILING DATE: 2002-05-13
PRIOR APPLICATION NUMBER: US 60/382, 409
PRIOR FILING DATE: 2002-05-21
PRIOR APPLICATION NUMBER: US 60/385, 280
PRIOR FILING DATE: 2002-06-03
PRIOR APPLICATION NUMBER: US 60/386, 879
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 64
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 25
LENGTH: 4234
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (863)...(2452)
US-11-312-958-25

Query Match 1.0%; Score 34; DB 7; Length 4234;
Best Local Similarity 53.8%; Pred. No. 11;
Matches 70; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

QY 2288 TGATTGGCATTGACAACTTGTGTGGTTCGGGAAAGGAGTTCATCCGTTGGGACGCTCA 2347
Db 1908 TGTTCGCGATCTTCAAGCTCTCCGCCACACTCCAAGGGGCTGACAGATCTGGGGCAAGACT 1967

QY 2348 GCAAGCTCTGGGGGAAAGGGGCTCCAGACAGCGCGATGTTTCTCTGTCAACAGCGTCCGC 2407
Db 1968 TGCAGGCGCTCATAGGAGGAGCTGGGGCTGCTCATCTTCTCTTCAACGGGCTCATCC 2027

QY 2408 TATACAGAG 2417
Db 2028 TCTTCTCAG 2037

RESULT 13
US-10-505-928-409
Sequence 409, Application US/10505928
Publication No. US20060088532A1
GENERAL INFORMATION:
APPLICANT: Ludwig Institute for Cancer Research et al.
TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
FILE REFERENCE: 28967/39178
CURRENT APPLICATION NUMBER: US/10/505, 928
CURRENT FILING DATE: 2004-08-27
PRIOR APPLICATION NUMBER: US 60/363, 019

```

;; PRIOR FILING DATE: 2002-03-07
;; NUMBER OF SEQ ID NOS: 866
;; SOFTWARE: PatentIn 3.2
;; SEQ ID NO: 409
;; LENGTH: 4573
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-10-505-928-409

Query Match 1.0%; Score 34; DB 6; Length 4573;
Best Local Similarity 59.2%; Pred. No. 11;
Matches 58; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

QY 1542 GTCTCCCACTGAGCCCGGACGAGGAGGCTCTCCCTTGATCAGCCCGCTGTAA 1601
DB 437 GTCCGGAGTTGGAGAGGAGAGCTCCAGAGATTTCGCCCCACCATCTCTCAGAGCTGGA 496

QY 1602 TGACCAAGGCTGCCCCGAGCAGAGATGAGATGAGG 1639
DB 497 TTCCCGGGCTGCACCTCGAGAACGAGAAAAAGAGG 534

RESULT 14
US-10-505-928-47

;; Sequence 47, Application US/10505928
;; Publication No. US20060088532A1

;; GENERAL INFORMATION:

;; APPLICANT: Ludwig Institute for Cancer Research et al.

;; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES

;; FILE REFERENCE: 28967/39178

;; CURRENT APPLICATION NUMBER: US/10/505,928

;; CURRENT FILING DATE: 2004-08-27

;; PRIOR APPLICATION NUMBER: US 60/363,019

;; PRIOR FILING DATE: 2002-03-07

;; NUMBER OF SEQ ID NOS: 866

;; SOFTWARE: PatentIn 3.2

;; SEQ ID NO 47

;; LENGTH: 4585

;; TYPE: DNA

;; ORGANISM: Homo sapiens

;; FEATURE:

;; OTHER INFORMATION: KIAA0062

US-10-505-928-47

Query Match 1.0%; Score 34; DB 6; Length 4585;
Best Local Similarity 59.2%; Pred. No. 11;
Matches 58; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

QY 1542 GTCTCCCACTGAGCCCGGACGAGGAGGCTCTCCCTTGATCAGCCCGCTGTAA 1601
DB 449 GTCCGGAGTTGGAGAGGAGAGCTCCAGAGATTTCGCCCCACCATCTCTCAGAGCTGGA 508

QY 1602 TGACCAAGGCTGCCCCGAGCAGAGATGAGATGAGG 1639
DB 509 TTCCCGGGCTGCACCTCGAGAACGAGAAAAAGAGG 546

QY 1602 TGACCAAGGCTGCCCCGAGCAGAGATGAGATGAGG 1639
DB 509 TTCCCGGGCTGCACCTCGAGAACGAGAAAAAGAGG 546

RESULT 15
US-11-217-529-174434

;; Sequence 174434, Application US/11217529

;; Publication No. US20060099612A1

;; GENERAL INFORMATION:

;; APPLICANT: SUNTORY LIMITED

;; APPLICANT: NAKAO, YOSHIHIRO

;; APPLICANT: NAKAMURA, NORIHISA

;; APPLICANT: KODAMA, YUKIKO

;; APPLICANT: FUJIMURA, TOMOKO

;; APPLICANT: ASHIKARI, TOSHIHIKO

;; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS

;; FILE REFERENCE: S-38-285

;; CURRENT APPLICATION NUMBER: US/11/217,529

;; CURRENT FILING DATE: 2005-09-02

;; PRIOR APPLICATION NUMBER: US 10/932,182

;; PRIOR FILING DATE: 2004-09-02
;; NUMBER OF SEQ ID NOS: 197023
;; SOFTWARE: PatentIn version 3.3
;; SEQ ID NO: 174434
;; LENGTH: 555
;; TYPE: DNA
;; ORGANISM: Saccharomyces pastorianus
US-11-217-529-174434

Query Match 1.0%; Score 33.8; DB 7; Length 555;
Best Local Similarity 48.2%; Pred. No. 3.7;
Matches 95; Conservative 0; Mismatches 102; Indels 0; Gaps 0;

QY 2580 GGAGAAAGTGGTTGAGGACATCCAGATGAGCCATTGACCTGGCGGAGAAAGCAGAGCCCC 2639
DB 246 GCACAGCTGTTTCTGGACAGAGAAAGCCATTAGATCTGAGCGAGAGCCCCGACACAC 305

QY 2640 CGCCCTGAGTTCCTGGCCAGAGCCCTTGACAAAGTCCCTGATGAAGCAACGCGC 2699
DB 306 AAACACTAGATCAATTAATTAATTAACAAATGACAGAGCCGCTTGACGAGAGACGCCGA 365

QY 2700 GGCTACCAAGAGTGAAGATGATGACCTGAGCGCCCTGGCAATGCTGAGAGCGGCAAGGC 2759
DB 366 CGACGACATTAATTATTATCAAGAGAGAGTCCATACATTTCTGCGCCGTGATTCTGCGC 425

QY 2760 CCGGACCGCGGACACA 2776
DB 426 CAGCAGCAGTAGCAACA 442

Search completed: May 22, 2006, 11:25:34
Job time : 36.587 secs


```
Db 668 TATGTTCAAGCTGAGCTTGATATACGATCTTAATGAGCATCTGAGAAATTACATCAG 727
Qy 586 TTAGCAAAAAATTAATACATACCTCTAGCAAGAGCATAGAGACAAATTCGGAAAT-T 644
Db 728 TGAATTTGAGATTTTCCCAAGAGATCAGCAAGATGAGAAAGAAATATGTGAAATTTCA 787
Qy 645 TCACCATTAACCATTCATGAGCAAAACACAGAGAAATGATTTCCAGCTCTAGAGATTGC 704
Db 788 TAAATAATGAACTCAGGGGCGAGGCCACCAAGTTGCTGAATTTAACTTCTCCTGAAAC 847
Qy 705 CCGTCGGCTAGAGATGTATGAAATCCGTTTGACCCCGGCCAAGAGACAGGAAAGCGAA 764
Db 848 TCACACTTTGGAAACCTACGGGGGTGAGTCTCACCATGCAAGATTCACAGGCGCAAC 907
Qy 765 GATCATCTGCGCGCTTCCCAACACGGGAATTCATAGTTTTCAGGGTTTCACTAAGATCA 824
Db 908 AACATTTTAAAGATTTACAGCTGACAGCTGAGCTTTGTGCTTTCAAGGAAATTAAGAAATCA 967
Qy 825 TGCCTTCAACTGGGCGCAAGGTGCGAAGCTGAGCTTCAAGAGAAACGCTTTCTCATCA 884
Db 968 TTGATTAATAATGGCCAGATGTCTGCBAATTTGAAGTGAAGAGAACATTTTATGTAT 1027
Qy 885 GCTCGGCGCAAGATGCCAATATGTCGTACCAAGATACCTTGGAAATTCCTAATGCGCAGTC 944
Db 1028 TGGCA-----CCCAAGAGAGAAAGAACCATGTTGGCATCAATATCTTCAACACC 1078
Qy 945 GGAATTTCTGCAAGTCTTCTGCAAAATCTGTGTTGAACATCATGCTTCTTTA 997
Db 1079 AGCTGCTGCACAAATCTTTTGAAGTGTGAGTGAAGAAACAGGCTTTTATA 1131
```

RESULT 2

```
US-11-412-475-5
; Sequence 5 Application US/11412475
; GENERAL INFORMATION:
; APPLICANT: The Salk Institute for Biological Studies
; APPLICANT: Hovatta, Ilirio
; APPLICANT: Verma, Indar M.
; APPLICANT: Lockhart, David J.
; APPLICANT: Barlow, Carolee
; TITLE OF INVENTION: ANXIETY RELATED FACTORS AND METHODS OF USE
; FILE REFERENCE: 7518-70020-02
; CURRENT APPLICATION NUMBER: US/11/412,475
; PRIOR FILING DATE: 2006-04-26
; PRIOR APPLICATION NUMBER: US 60/675,604
; PRIOR FILING DATE: 2005-04-27
; NUMBER OF SEQ ID NOS: 115
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 5
; LENGTH: 2487
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (52) .. (1716)
US-11-412-475-5
```

Query Match 2.9%; Score 99; DB 7; Length 2487;

Best Local Similarity 48.5%; Pred. No. 4,4e-16;

Matches 378; Conservative 0; Mismatches 380; Indels 21; Gaps 3;

```
Qy 238 GTGCTGCTGATGATGAGTTTGAACACCTCAACCTGTGGAAGTACTATTTTGGCCCTC 297
Db 169 GTTGCTCTTGACCACTTTTCCGTCAATAACTGTGGAATATTTTGGGCTG 228
Qy 298 GAGTTTCTGATCAAAAAAGATCAGGTGTGCTGGAATCTCTAAACCATTTGTGAA 357
Db 229 CCGTACTGTGACAGAAAGCATCAGATATTTGGCTGATCCGCAAAAACCTTGACAGAA 288
Qy 358 CA-----GATTGAAGGCCAAAGCAAGTGTGTTGAAGTTGTGTTGTTGTTGTTT 408
Db 289 CACAGAGAGCTATCAACACTGGAAGCTCAATATCTTGTATTTTGGATTAATTTCTAT 348
```

```
Qy 409 CCGCTGACACACACAACCTCCGAAGAACTCACAAAGTACCTGTGCGCTGCAAGTG 468
Db 349 GGTGAAGATTCATGTAACTCAAGAAAGAAATTAACAGATATCAATTTTCTTACAGTG 408
Qy 469 AAGCAAGACTTGAGCTCAAGGAGTTGACGTGTAAATGACACCAAGGAGCTCTTTGAT 528
Db 409 AAGCAAGAGTCCCTTCAAGGCGGCTGCTTCCCTGTCAATATTTGCTGATGAGAGG 468
Qy 529 TCACACTTGTGCAATCTGAGATTGGGATTTTGTATGAAGCTTGGACAGAGACCTTA 588
Db 469 GATATGCTCATCAGGCGCAACTGGAAGCACAGACCCATTAACACACACTGCGGGTAT 528
Qy 589 GCAAAAAATTAATATAT-----ACCTGACAGAGAGCACTAGAGACAAATTCGTGAA 642
Db 529 GTGTGGAGTACCGGTTTGTCTTCTATCAGAAAGAGAGCTGGAAGAACATTAAGAG 588
Qy 643 TTTCACATTAACCATTTGAGACAAACACCAAGATCAGATTTCCAGCTCTTAAGATT 702
Db 589 ATTCATTAATACTCTAATGGGTGAGGCTCTTCCGAAGCTGAGCTGAATTAATTTGAGAG 648
Qy 703 GCCCGTGGCTAGAGATGTATGGAATCCGTTGCAACCGGCGCAAGAGACAGGAGCGAC 762
Db 649 GCAAAATCCCTGAGAGATGTATGTGTGACCTTCATCTGTATGAGAGAAATTAAGTTC 708
Qy 763 AAGATCAATCTGCGCGTTGCCAACACGGGAATTCATGTTTCAAGGTTTCACTAAGATC 822
Db 709 GAGTACTTCTTAAGGCTTGATCTCATCGGGAAGTGTGTCTAATAAATTAAGAAAGT 768
Qy 823 AATGCTTCAACTGGGCGCAAGGTGCGAAGCTGAGCTTCAGAGAGAGAGAGCGCTTTCATC 882
Db 769 GGAAGTATTTCTGCGCTCGATTAACAAAGTGACATTCAGAGAAACCAAGCTTGAAGCTC 828
Qy 883 AAGCTCGGCGAGATGCCAATATGTGCTACAGATATCTTGAATTCCTGATGGCCAGT 942
Db 829 AAGTCTTGGGAAAGACTGTATATAAATCTCATCTTTTGAAGCTGAAAGCAAACT 888
Qy 943 CCGGATTTCTGCAAGTCTTCTGGAATAATCTGTGTTGAACATCATGCTTCTTTAGACT 1001
Db 889 GCT-----TGCACACACTCTGGAATGACGCGTGAAGCACATATGCTTTTCAAGAT 941
```

RESULT 3

```
US-11-412-475-7
; Sequence 7 Application US/11412475
; GENERAL INFORMATION:
; APPLICANT: The Salk Institute for Biological Studies
; APPLICANT: Hovatta, Ilirio
; APPLICANT: Verma, Indar M.
; APPLICANT: Lockhart, David J.
; APPLICANT: Barlow, Carolee
; TITLE OF INVENTION: ANXIETY RELATED FACTORS AND METHODS OF USE
; FILE REFERENCE: 7518-70020-02
; CURRENT APPLICATION NUMBER: US/11/412,475
; PRIOR FILING DATE: 2006-04-26
; PRIOR APPLICATION NUMBER: US 60/675,604
; PRIOR FILING DATE: 2005-04-27
; NUMBER OF SEQ ID NOS: 115
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 7
; LENGTH: 2332
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (275) .. (2026)
US-11-412-475-7
```

Query Match 2.9%; Score 98.4; DB 7; Length 2332;

Best Local Similarity 48.4%; Pred. No. 6,1e-16;

Matches 381; Conservative 0; Mismatches 386; Indels 21; Gaps 3;

```
Qy 238 GTGCTGCTGATGATGAGTTTGAACACCTCAACCTGTGGAAGTACTATTTTGGCCCTC 297
```

```
Db 392 GTTGCTCTGACACGATATTCATCACTGAACCTTGATGAGATGATTTATTTTGGGCTA 451
Qy 298 GAGTTCTCGATACACAAAAGATCAGGATGCTGATCTCTAAACCATTGGAAA 357
Db 452 CGTTACTGTGACGAAGCCATCAGATATTTGGCTGATCTCTCAAAAACCTCTGTGAA 511
Qy 358 CA-----GATTGAAGGCCAAGCAGCTGTTGTTAAGTTGGGGAATCTTT 408
Db 512 CACAAAGAACTGATCAACTGAGACCTCCATATCTTTGTTTGGTATTAATTTAT 571
Qy 409 CCGCTTGACACACACAACTCCAAAGAAAGTCAAGATCCTGTTCCGCTGAGGATG 468
Db 572 GCGAAGATCCATGTAACCTTAAGAAATAACAGATATAGTTTTTTCTTGCAAGTG 631
Qy 469 AAGCAGACTTGCTCAAGCAGCTTGACGTGATGACCAAGCCAGCTCTCTTGANT 528
Db 632 AAGCAAGATGTCCTTCAAGGCGCTGCTGCCGTCAACCTGCTCAGCTGAGGGA 691
Qy 529 TCACACATTGTGCACTGATGATTTGGGATTTTG-----ATGAGCCTTGACAGAG 582
Db 692 GCGTATCCATCCAGTCGAGCTTGAGATTAACCATATTAACATCTGAGAGAT 751
Qy 583 CACTTAGCAAAAATAATATACATCTCAAGCAAGCAGCTAGAGCAAAATCGTGAA 642
Db 752 GATCTGAGTACCGGTTTGTCTGATCAGAGAAAGAACTTGAAGAGCATTAGAAAG 811
Qy 643 TTTCACCAATACCACTTGGAACAACAACAGAAATCAGATTTCCAGCTCTAGAGAT 702
Db 812 ATTCAATAAACTTAATGGGTGAGATTTCTTCTGAGGCTGAGCTGATTAATTGAGGACT 871
Qy 703 GCCCGTGGCTAGATGATGGAATCCGTTGACCCCGCCAAAGACAGGGAAGCAGC 762
Db 872 GCCAAATCCCTGAGAGATGATGCGCTTACCTCCGTCATGAGAGAAACAAAGTCT 931
Qy 763 AAGATCAATCTGCGCTGCGCAACAGGAAATCTAGTGTTCAGGATTTCACTAAGATC 822
Db 932 GAGTATTTCTGAGATTAATCTCGGTTGTGTGTGTGTAAGAAATAAAGCAAGTG 991
Qy 823 AATGCTTCACTGGGCGCAAGTGCGAGAGCTGAGCTTCAAGAGAAAGCGCTTTCATC 882
Db 992 GGGAAATATTTCTGGGCTCGGATTAACAAGGTTCACTCAAGGAGACTCAATTTGAATC 1051
Qy 883 AAGCTCGGCGCAGATCCAAATAGTGCTCCAGAGATACCTTGGAATTCCTGATGGCAGT 942
Db 1052 AAGATCTGAGAAAGATTTGTAAGAACTCACTCTTTTGAAGCTCGG-----AGT 1105
Qy 943 CCGGATTTCTGCAAGTCTTTCTGAAATCTGTGTAACATCATGCTCTTTAGACTT 1002
Db 1106 AAAACTGCTTGCAAGACCTCTGGAAGTGAGTGGAACATCATATTTTAAAGATG 1165
Qy 1003 TTGAAGA 1010
Db 1166 CCAGAAAA 1173

RESULT 4
US-11-418-718-2805
; Sequence 2805, Application US/11418718
; GENERAL INFORMATION:
; APPLICANT: ROSETTA GENOMICS LTD
; APPLICANT: Benitich, Itzhak
; APPLICANT: Avniel, Amir
; APPLICANT: Karov, Yael
; TITLE OF INVENTION: DIFFERENTIALLY EXPRESSED MICRORNA IN LIVER LUNG AND BLUDDER CANCER
; FILE REFERENCE: Efs-56
; CURRENT FILING DATE: 2006-05-04
; NUMBER OF SEQ ID NOS: 8661
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2805
; LENGTH: 63
```

```
; TYPE: RNA
; ORGANISM: Human
US-11-418-718-2805
Query Match 1.6%; Score 56; DB 7; Length 63;
Best Local Similarity 67.9%; Pred. No. 1e-05;
Matches 38; Conservative 18; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 3387 GGCATCCGCTGGGGGCGCTGTTCTTACGTAAGCCAGATTAATAACATTCGATT 3442
Db 1 GGCATCCGCTGGGGGCGCTGTTCTTACGTAAGCCAGATTAATAACATTCGATT 3442

RESULT 5
US-11-418-870-2805
; Sequence 2805, Application US/1141870
; GENERAL INFORMATION:
; APPLICANT: ROSETTA GENOMICS LTD
; APPLICANT: Benitich, Itzhak
; APPLICANT: Avniel, Amir
; APPLICANT: Karov, Yael
; TITLE OF INVENTION: DIFFERENTIALLY EXPRESSED MICRORNA IN LIVER LUNG AND BLUDDER CANCER
; FILE REFERENCE: Efs-56
; CURRENT FILING DATE: 2006-05-04
; NUMBER OF SEQ ID NOS: 8661
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2805
; LENGTH: 63
; TYPE: RNA
; ORGANISM: Human
US-11-418-870-2805
```

```
Query Match 1.6%; Score 56; DB 7; Length 63;
Best Local Similarity 67.9%; Pred. No. 1e-05;
Matches 38; Conservative 18; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 3387 GGCATCCGCTGGGGGCGCTGTTCTTACGTAAGCCAGATTAATAACATTCGATT 3442
Db 1 GGCATCCGCTGGGGGCGCTGTTCTTACGTAAGCCAGATTAATAACATTCGATT 3442

RESULT 6
US-11-418-875-2805
; Sequence 2805, Application US/1141875
; GENERAL INFORMATION:
; APPLICANT: ROSETTA GENOMICS LTD
; APPLICANT: Benitich, Itzhak
; APPLICANT: Avniel, Amir
; APPLICANT: Karov, Yael
; TITLE OF INVENTION: DIFFERENTIALLY EXPRESSED MICRORNA IN LIVER LUNG AND BLUDDER CANCER
; FILE REFERENCE: Efs-56
; CURRENT FILING DATE: 2006-05-04
; NUMBER OF SEQ ID NOS: 8661
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2805
; LENGTH: 63
; TYPE: RNA
; ORGANISM: Human
US-11-418-875-2805
```

```
Query Match 1.6%; Score 56; DB 7; Length 63;
Best Local Similarity 67.9%; Pred. No. 1e-05;
Matches 38; Conservative 18; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 3387 GGCATCCGCTGGGGGCGCTGTTCTTACGTAAGCCAGATTAATAACATTCGATT 3442
Db 1 GGCATCCGCTGGGGGCGCTGTTCTTACGTAAGCCAGATTAATAACATTCGATT 3442
```



```
RESULT 7
US-11-418-718-5577
; Sequence 5577, Application US/11418718
; GENERAL INFORMATION:
; APPLICANT: ROSETTA GENOMICS LTD
; APPLICANT: Benlwich, Itzhak
; APPLICANT: Avnlel, Amir
; APPLICANT: Karov, Yael
; APPLICANT: Aharonov, Rantl
; TITLE OF INVENTION: DIFFERENTIALLY EXPRESSED MICRORNA IN LIVER LUNG AND BLUDDER CANCER
; FILE REFERENCE: Efs-56
; CURRENT APPLICATION NUMBER: US/11/418,718
; CURRENT FILING DATE: 2006-05-04
; NUMBER OF SEQ ID NOS: 8661
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 5577
; LENGTH: 62
; TYPE: RNA
; ORGANISM: Human
US-11-418-718-5577
```

```
Query Match 1.6%; Score 55; DB 7; Length 62;
Best Local Similarity 67.3%; Pred. No. 1.9e-05;
Matches 37; Conservative 18; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 3388 GCATCCGCTGGGGGCGCTGTTCTTAGCTAGTGCAGATTAATAACATTGTCATT 3442
DB 1 GCAUCCGCGUGGGGCGCGCUGUCUUAAGCUAGUCCAGUAAUUAACAUUGUCAUU 55
```

```
RESULT 8
US-11-418-870-5577
; Sequence 5577, Application US/11418870
; GENERAL INFORMATION:
; APPLICANT: ROSETTA GENOMICS LTD
; APPLICANT: Benlwich, Itzhak
; APPLICANT: Avnlel, Amir
; APPLICANT: Karov, Yael
; APPLICANT: Aharonov, Rantl
; TITLE OF INVENTION: DIFFERENTIALLY EXPRESSED MICRORNA IN LIVER LUNG AND BLUDDER CANCER
; FILE REFERENCE: Efs-56
; CURRENT APPLICATION NUMBER: US/11/418,870
; CURRENT FILING DATE: 2006-05-04
; NUMBER OF SEQ ID NOS: 8661
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 5577
; LENGTH: 62
; TYPE: RNA
; ORGANISM: Human
US-11-418-870-5577
```

```
Query Match 1.6%; Score 55; DB 7; Length 62;
Best Local Similarity 67.3%; Pred. No. 1.9e-05;
Matches 37; Conservative 18; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 3388 GCATCCGCTGGGGGCGCTGTTCTTAGCTAGTGCAGATTAATAACATTGTCATT 3442
DB 1 GCAUCCGCGUGGGGCGCGCUGUCUUAAGCUAGUCCAGUAAUUAACAUUGUCAUU 55
```

```
RESULT 9
US-11-429-720-2923
; Sequence 2923, Application US/11429720
; GENERAL INFORMATION:
; APPLICANT: ROSETTA GENOMICS LTD
; APPLICANT: Benlwich, Itzhak
; APPLICANT: Avnlel, Amir
; APPLICANT: Karov, Yael
; APPLICANT: Aharonov, Rantl
; TITLE OF INVENTION: DIFFERENTIALLY EXPRESSED MICRORNA IN PROSTATE CANCER
; FILE REFERENCE: Efs-59
; CURRENT APPLICATION NUMBER: US/11/429,720
; CURRENT FILING DATE: 2006-05-08
```

```
; NUMBER OF SEQ ID NOS: 3751
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2923
; LENGTH: 62
; TYPE: RNA
; ORGANISM: Homo sapiens
US-11-429-720-2923
```

```
Query Match 1.6%; Score 55; DB 7; Length 62;
Best Local Similarity 67.3%; Pred. No. 1.9e-05;
Matches 37; Conservative 18; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 3388 GCATCCGCTGGGGGCGCTGTTCTTAGCTAGTGCAGATTAATAACATTGTCATT 3442
DB 1 GCAUCCGCGUGGGGCGCGCUGUCUUAAGCUAGUCCAGUAAUUAACAUUGUCAUU 55
```

```
RESULT 10
US-11-418-875-5577
; Sequence 5577, Application US/11418875
; GENERAL INFORMATION:
; APPLICANT: ROSETTA GENOMICS LTD
; APPLICANT: Benlwich, Itzhak
; APPLICANT: Avnlel, Amir
; APPLICANT: Karov, Yael
; APPLICANT: Aharonov, Rantl
; TITLE OF INVENTION: DIFFERENTIALLY EXPRESSED MICRORNA IN LIVER LUNG AND BLUDDER CANCER
; FILE REFERENCE: Efs-56
; CURRENT APPLICATION NUMBER: US/11/418,875
; CURRENT FILING DATE: 2006-05-04
; NUMBER OF SEQ ID NOS: 8661
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 5577
; LENGTH: 62
; TYPE: RNA
; ORGANISM: Human
US-11-418-875-5577
```

```
Query Match 1.6%; Score 55; DB 7; Length 62;
Best Local Similarity 67.3%; Pred. No. 1.9e-05;
Matches 37; Conservative 18; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 3388 GCATCCGCTGGGGGCGCTGTTCTTAGCTAGTGCAGATTAATAACATTGTCATT 3442
DB 1 GCAUCCGCGUGGGGCGCGCUGUCUUAAGCUAGUCCAGUAAUUAACAUUGUCAUU 55
```

```
RESULT 11
US-11-431-708-6398
; Sequence 6398, Application US/11431708
; GENERAL INFORMATION:
; APPLICANT: RUBEN, Steven
; TITLE OF INVENTION: BREAST DISEASE TARGETS AND USES THEREOF
; FILE REFERENCE: CLO01599-ORD
; CURRENT APPLICATION NUMBER: US/11/431,708
; CURRENT FILING DATE: 2006-05-11
; NUMBER OF SEQ ID NOS: 8544
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6398
; LENGTH: 5857
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-431-708-6398
```

```
Query Match 1.2%; Score 40.4; DB 7; Length 5857;
Best Local Similarity 46.1%; Pred. No. 0.66;
Matches 177; Conservative 0; Mismatches 201; Indels 6; Gaps 1;
```

```
OY 2059 CAGAAGTGTGTGTTACTACCGCTCAACACCTTCCTCGGCGCACTGCACCGGCTCATG 2118
DB 564 CGAAGACACACGAGACATCCCTTTGGAAGGCTACCTGTTGTCTCCGATCCAGAGATCTGC 623
OY 2119 CACTACAAGCAGGTCCTGAGCGGCTGTGCAAAACACACCGCGAGCCAGCGCACTTC 2178
```

Position	Sequence	Position	Sequence
Db	624 AAGTACC CGCTCTCTTAAAGAGTGGCCAAAGAGACTCCCGCAGACCCAGACCAC	683	
Qy	2179 AGGCACTGCCAGGCGCTTTTGGCAGAGATCACGAGATGATGGACAGCTCCACGTTACG	2238	
Db	684 CCGGGGTCCAGAGTGCCCTCAGGCCATGAAACCGTTTCTCCAACTCAATGAAGACC	743	
Qy	2239 ATGATCAAGATGAGAAATTTCCAGAACTGCACGAACTCAAGAAAGATTGATTTGGCATTT	2298	
Db	744 AAGCGGCAGATGAGAAAGCTGAAAGCCCTTGAGCAGCTGCAGTCCCATGAAAGGCTGG	803	
Qy	2299 GA-----CAATCTTGTGTTCCGGGAAGGAGTTCAATCCGCTTGGGACGCTCAGCAAG	2352	
Db	804 GAGGGTTCAACTCACAGACATCTGCACCTAGCTCTCTCTGGCAAGGACCTTTGTAAAG	863	
Qy	2353 CTCTCGGGGAAGGGGCTTCAGCAGCGCATGTCTTCTCTGTTCAACGACGTCTGCTATAC	2412	
Db	864 ATCTCTGGGGCAACATCCAGAAAGGCGCTTCTCTCTTCGACAACTTCTCGTAC	923	
Qy	2413 ACGAGCGGGGGCTGAGGCGCTCC	2436	
Db	924 TGCAAGCGAAATCCAGGGTACC	947	

```

RESULT 12
US-10-669-920-1193
/ Sequence 1193, Application US/10669920
/ GENERAL INFORMATION:
/ APPLICANT: Morris, David W.
/ APPLICANT: Malandro, Marc S.
/ TITLE OF INVENTION: NOVEL THERAPEUTIC TARGETS IN CANCER
/ FILE REFERENCE: 20366-066001
/ CURRENT APPLICATION NUMBER: US/10/669,920
/ CURRENT FILING DATE: 2003-09-23
/ PRIOR APPLICATION NUMBER: US 10/004,113
/ PRIOR FILING DATE: 2001-10-23
/ PRIOR APPLICATION NUMBER: US 10/052,482
/ PRIOR FILING DATE: 2001-11-08
/ PRIOR APPLICATION NUMBER: US 09/997,722
/ PRIOR FILING DATE: 2001-11-30
/ PRIOR APPLICATION NUMBER: US 10/034,650
/ PRIOR FILING DATE: 2001-12-20
/ PRIOR APPLICATION NUMBER: US 10/085,117
/ PRIOR FILING DATE: 2002-02-27
/ PRIOR APPLICATION NUMBER: US 10/087,192
/ PRIOR FILING DATE: 2002-03-01
/ PRIOR APPLICATION NUMBER: US 10/322,281
/ PRIOR FILING DATE: 2002-12-17
/ PRIOR APPLICATION NUMBER: US 10/322,696
/ PRIOR FILING DATE: 2002-12-17
/ NUMBER OF SEQ ID NOS: 1441
/ SOFTWARE: FASTSEQ for Windows Version 4.0
/ SEQ ID NO 1193
/ LENGTH: 63810
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (1)...(63810)
/ OTHER INFORMATION: n = A,T,C or G
/ US-10-669-920-1193

```

	Query Match	Similarity	1.1%;	Score 38.8;	DB 6;	Length 63810;
	Best Local	Similarity	49.0%;	Prod. No. 4.5;		
	Matches	103;	Conservative	0;	Mismatches	107;
				Indels	0;	Gaps
						0;
QY	2460	GCTCCCGCTCTATGCGATGACGATTAAGAGACGAGACGATGCGGGGCTGCCCACTG				2519
DB	30746	GCCCCAGATGCTGTATGCTATGCTGAACCTCGACGCGGACAGCCCTATAGACATCCCGCGCT				30805
QY	2520	CTTGACCTCTCCGGGGCCAGCGGACGATGCATCATCTGTGGCCGCGCAGTTCCTGGTCCGAGAT				2579
DB	30806	GCTGTCTCTTACAGTCGCGAGCTGAAGGCGCGGACGCTGACCATGTGGGCTCGGTGCT				30865

QY	2580	GGAGAAATGGGTGGTGGAGCAATCCAGANNGCATGACCTGGGGGGAGAAAGACGAGGCC	2639
Db	30866	GGAGGGGACCTTACTCTGACCAAGACACATGAGAGCTCAGCGGGCCGAGAGAGTGGGCCCGGAC	30925
QY	2640	CGCCCCCTGAAGTTCCTGGCCAGCAGCCCCCCC	2669
Db	30926	GGGGTCTGGGGGCGCAGGCTCTTCTCCACAC	30955

```

RESULT 13
US-10-461-194A-2
; Sequence 2, Application US/10461194A
; GENERAL INFORMATION:
; APPLICANT: Hutchinson, Richard C.
; APPLICANT: Reid, Ralph C.
; APPLICANT: Hu, Zhihao
; APPLICANT: Rascher, Andreas
; APPLICANT: Schimer, Andreas
; APPLICANT: McDaniel, Robert
; TITLE OF INVENTION: RECOMBINANT POLYNUCLEOTIDES ENCODING
; TITLE OF INVENTION: PRO-GLUTAMINACIN PRODUCING POLYPEPTIDE SYNTHASES AND
; TITLE OF INVENTION: ACCESSORY PROTEINS, AND USES THEREOF
; FILE REFERENCE: 30062009700
; CURRENT APPLICATION NUMBER: US/10/461,194A
; CURRENT FILING DATE: 2003-06-13
; PRIOR APPLICATION NUMBER: US 60/389,255
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: US 60/393,929
; PRIOR FILING DATE: 2002-07-03
; PRIOR APPLICATION NUMBER: US 60/395,275
; PRIOR FILING DATE: 2002-07-12
; PRIOR APPLICATION NUMBER: US 10/212,962
; PRIOR FILING DATE: 2002-08-05
; PRIOR APPLICATION NUMBER: US 60/415,326
; PRIOR FILING DATE: 2002-09-30
; PRIOR APPLICATION NUMBER: US 60/420,820
; PRIOR FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: US 60/433,130
; PRIOR FILING DATE: 2002-12-13
; NUMBER OF SEQ ID NOS: 161
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 86941
; TYPE: DNA
; ORGANISM: Streptomyces hygroscopicus
US-10-461-194A-2

```

Query Match	1.1%	Score 38.6	DB 6	Length 86941
Best Local Similarity	42.0%	Pred. No. 5.8	Matches 224	Conservative 0
			Mismatches 309	Indels 0
			Gaps 0	
Qy	2322	GGAATTATCATCGCTCGGGCAGCCTCAGCAAGCTCTCGGGAAAGGGCTCCACAGCGCAT	2381	
Db	61350	GGAGGCGGAAACGGTTGGCATTCACCCCGCGCTGCTGAAAGCCCGCTCACAAGCGGCGAG	62009	
Qy	2382	GTTCTTCCTGTTCAACGACGCTCTGTTATACAGAGCCGGGGCTGACCGGCTCCATCA	2441	
Db	62010	CTTCTGCTGCTCCCTCGGAGCCCGGCGGACAGGATCACTTGCTGCGCTTGGAACA	62069	
Qy	2442	GTTTAAAGTCCACGGGACGCTCCCGCTCTATGGCATGACGATTGAGGAAGCGAAGACGA	2501	
Db	62070	CGTCCGTCTGCACGCGGGCGGCGCTCGCGGCTCGCGGTGCATGTCTCCGCCGCTGGCGCA	62129	
Qy	2502	GTTGGGGGTGCCCCACATGCCCTACCTCCGGGGGCGACGGGCGATCATCATGTGGGCGC	2561	
Db	62130	CGAGCGCTTCTCGGTACGTGTGCCAGCGACTTCGGGCCAGACGGTGGCTTCCGTGATCTC	62189	
Qy	2562	CAGTTCTCGGTCGAGATGGAAGATGGGTTTGAGCATTCAGANTGGCCATTGACCTGGC	2621	
Db	62190	GCTCACCTCTCGCGCGGTGTATCCGGCCCAACTGAGATTCGGTACGGCCGAGACGAGCGGT	62249	
Qy	2622	GGAAGAAAGCAGACGCCCGCCCTTGATGTTCTTGCGCAGACGCCCCCTTGACACAAAGTC	2681	

Db 62250 GTGACGGTCCCTGGAGGAGACCTCCTGCAGGCGCGATCTCTGGGCGCGCT 62309
Qy 2682 CCTATGAAGCACCAGCGCTGACCAAGATCAAGATGACTGAGGCGCTCGGCAC 2741
Db 62310 GGGCGATCCGGCCACCGTGGCGCCGGGGCCATGGAGACCCGAAAGCGGGAGCGC 62369
Qy 2742 ATCGCTGAGCGCCAGCGCCCGCAGCCGCAACAAATGTGACGTGTGCTGCGACCG 2801
Db 62370 CCTCCCGACGTCCTGTGTGGCCGATACGGCGCGCTGGCGCAAGACCTCACCGGCGCGC 62429
Qy 2802 CAACACGACGTCCTCATGATGAGCTTCAGCATGCGATGGAGATCACTTGT 2854
Db 62430 GACCGCGCGGGCCCGCGGCTCACCGCGCGCTGTGGCGGAGATCAAGCGGT 62482

RESULT 14
US-10-669-920-398/c
; Sequence 398, Application US/10669920
; GENERAL INFORMATION:
; APPLICANT: Moris, David W.
; APPLICANT: Malandro, Marc S.
; TITLE OF INVENTION: NOVEL THERAPEUTIC TARGETS IN CANCER
; FILE REFERENCE: 20366-066001
; CURRENT APPLICATION NUMBER: US/10/669,920
; PRIOR FILING DATE: 2003-09-23
; PRIOR APPLICATION NUMBER: US 10/004,113
; PRIOR FILING DATE: 2001-10-23
; PRIOR APPLICATION NUMBER: US 10/052,482
; PRIOR FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: US 09/997,722
; PRIOR FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: US 10/034,650
; PRIOR FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: US 10/085,117
; PRIOR FILING DATE: 2002-02-27
; PRIOR APPLICATION NUMBER: US 10/087,192
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 10/322,281
; PRIOR FILING DATE: 2002-12-17
; PRIOR APPLICATION NUMBER: US 10/322,696
; PRIOR FILING DATE: 2002-12-17
; NUMBER OF SEQ ID NOS: 1441
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 398
; LENGTH: 1394
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-669-920-398

Query Match 1.1%; Score 37.2; DB 6; Length 1394;
Best Local Similarity 43.1%; Pred. No. 2.4;
Matches 180; Conservative 0; Mismatches 238; Indels 0; Gaps 0;
Qy 2322 GGAGTTCATCCGTCTGGGAGCCTCAGCAAGCTTCGGGAAAGGGCTTCAGACGCAT 2381
Db 1120 GGGGCTCATGGAGGCGGGTTTCGACACTGTCTGGGAGATGAAGACCAAGAGGGGCG 1061
Qy 2382 GTTCTTCTGTTCAAGAGAGTCTCTGTATACAGAGAGCGGGGGCTGACGGCTCCAAATA 2441
Db 1060 CCAACACAGCTGGGCGCACCATGATCCCGAGAGCTAGCCATCATCTTCCACCTCGGA 1001
Qy 2442 GTTAAAGTCAAGGAGCAGCTCCCGCTTATGGCATGACGATTGAGAGAGCAAGACA 2501
Db 1000 GTCCCGGAGACGTGTGGGGCGCTCTCTGCGCGTGGTGAAGTGGACAAGGTGGCCAC 941
Qy 2502 GTGGGGGTGCCCCCACTGCTGACCTTCGGGGCCAGCGGCACTCATCTGTCGCGC 2561
Db 940 GGTGACCGTGTTCAGAGAGAGACAGCCGACCGAGAGGCGCGCCAGCATGAGAAAG 881
Qy 2562 CAGTTCTCGTTCGAGATGAGAAAGTGGTTGAGAGCATTCAGATGGCCATTGACCTGGC 2621
Db 2562 CAGTTCTCGTTCGAGATGAGAAAGTGGTTGAGAGCATTCAGATGGCCATTGACCTGGC 2621
Qy 880 CAGCCCGAAGGCACTGCCCGGACTCGCCAGCTGACGAGAAACACCAAGAACCCCGC 821

Qy 2622 GGAGAAGACAGACGCCCGCCCTGAGTTCTTGCCAGACAGCCCGCTGACAAAGTTC 2681
Db 820 GATTTGACAGGTGTAGCCACACCCAGACAGGGGCGAGCGCCACAGGCGCAAGCGCGC 761
Qy 2682 CCTATGAAGCACCAGCGCTGACCAAGATCAAGATGACTGAGGCGCTCGGC 2739
Db 760 CCGCCACACAGCCCGCCAGGCTGGCCAGGCGCGCTGCGAGGCGACCGCGCGCGC 703

RESULT 15
US-10-669-920-400/c
; Sequence 400, Application US/10669920
; GENERAL INFORMATION:
; APPLICANT: Moris, David W.
; APPLICANT: Malandro, Marc S.
; TITLE OF INVENTION: NOVEL THERAPEUTIC TARGETS IN CANCER
; FILE REFERENCE: 20366-066001
; CURRENT APPLICATION NUMBER: US/10/669,920
; PRIOR FILING DATE: 2003-09-23
; PRIOR APPLICATION NUMBER: US 10/004,113
; PRIOR FILING DATE: 2001-10-23
; PRIOR APPLICATION NUMBER: US 10/052,482
; PRIOR FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: US 09/997,722
; PRIOR FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: US 10/034,650
; PRIOR FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: US 10/085,117
; PRIOR FILING DATE: 2002-02-27
; PRIOR APPLICATION NUMBER: US 10/087,192
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 10/322,281
; PRIOR FILING DATE: 2002-12-17
; PRIOR APPLICATION NUMBER: US 10/322,696
; PRIOR FILING DATE: 2002-12-17
; NUMBER OF SEQ ID NOS: 1441
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 400
; LENGTH: 1507
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-669-920-400

Query Match 1.1%; Score 37.2; DB 6; Length 1507;
Best Local Similarity 43.1%; Pred. No. 2.5;
Matches 180; Conservative 0; Mismatches 238; Indels 0; Gaps 0;
Qy 2322 GGAGTTCATCCGTCTGGGAGCCTCAGCAAGCTTCGGGAAAGGGCTTCAGACGCAT 2381
Db 1120 GGGGCTCATGGAGGCGGGTTTCGACACTGTCTGGGAGATGAAGACCAAGAGGGGCG 1061
Qy 2382 GTTCTTCTGTTCAAGAGAGTCTCTGTATACAGAGAGCGGGGGCTGACGGCTCCAAATA 2441
Db 1060 CCAACACAGCTGGGCGCACCATGATCCCGAGAGCTGAGCCATCATCTTCCACCTCGGA 1001
Qy 2442 GTTAAAGTCAAGGAGCAGCTCCCGCTTATGGCATGACGATTGAGAGAGCAAGACA 2501
Db 1000 GTCCCGGAGACGTGTGGGGCGCTCTCTGCGCGTGGTGAAGTGGACAAGGTGGCCAC 941
Qy 2502 GTGGGGGTGCCCCCACTGCTGACCTTCGGGGCCAGCGGCACTCATCTGTCGCGC 2561
Db 940 GGTGACCGTGTTCAGAGAGAGACAGCCCGACCGAGAGGCGCGCCAGCATGAGAAAG 881
Qy 2562 CAGTTCTCGTTCGAGATGAGAAAGTGGTTGAGAGCATTCAGATGGCCATTGACCTGGC 2621
Db 880 CAGCCCGAAGGCACTGCCCGGACTCGCGCCAGCGCTGACGAGAAACACCAAGAACCCCGC 821
Qy 2622 GGAAGAAGACAGACGCCCGCCCTGAGTTCTTGCCAGACAGCCCGCTGACAAAGTTC 2681
Db 820 GATTTGACAGGTGTAGCCACACCCAGACAGGGGCGAGCGCCACAGGCGCAAGCGCGC 761
Qy 2682 CCTATGAAGCACCAGCGCTGACCAAGATCAAGATGACTGAGGCGCTCGGC 2739

This Page Blank (usrtr)


```
|||||
Db 229 CTTGGGAAAGTGCTGCTGATGCAAGTTTGCACCACTCAACCTCGTGGAAAGTGACTAT 288
|||
Qy 241 TTTGGCTGAGTTTCTCTATCAAAAAGATCAAGGTGTGGATGATCTCTAAACCC 300
|||
Db 289 TTTGGCTGAGTTTCTCTATCAAAAAGATCAAGGTGTGGATGATCTCTAAACCC 348
|||
Qy 301 ATTGGAACAGATTGAAGGCCAAAGCAGTGTGTGAATTGATTGTGTGAATTCTT 360
|||
Db 349 ATTGGAACAGATTGAAGGCCAAAGCAGTGTGTGAATTGATTGTGTGAATTCTT 408
|||
Qy 361 CGGCTGACCAACACCACTCCAAAGAACTTCAAGGTACTGTTCGCGCTGCAGGTG 420
|||
Db 409 CGGCTGACCAACACCACTCCAAAGAACTTCAAGGTACTGTTCGCGCTGCAGGTG 468
|||
Qy 421 AAGCAGGACTTGGCTCAAGGCAAGTTGAGCTGTAAAGACCAAGCCACTCTTTGAT 480
|||
Db 469 AAGCAGGACTTGGCTCAAGGCAAGTTGAGCTGTAAAGACCAAGCCACTCTTTGAT 528
|||
Qy 481 TCACACATTGCTCAATCTGAGATTGGGATTTTGTATGAAGCTTGGACAGAGCACTTA 540
|||
Db 529 TCACACATTGCTCAATCTGAGATTGGGATTTTGTATGAAGCTTGGACAGAGCACTTA 588
|||
Qy 541 GCAAAAATTAATACATACCTTACAGCAAGCCTAGAGACAAAATCGTGAATTTCAC 600
|||
Db 589 GCAAAAATTAATACATACCTTACAGCAAGCCTAGAGACAAAATCGTGAATTTCAC 648
|||
Qy 601 CATTAACCAATTGGACAAACACAGAGAAATGATTTCAAGTCTCTAAGATTGCCGT 660
|||
Db 649 CATTAACCAATTGGACAAACACAGAGAAATGATTTCAAGTCTCTAAGATTGCCGT 708
|||
Qy 661 CGGCTAGAGATGATGGAATCCGTTGCAACCGGCCAAAGACAGGAGGCAAGAGATC 720
|||
Db 709 CGGCTAGAGATGATGGAATCCGTTGCAACCGGCCAAAGACAGGAGGCAAGAGATC 768
|||
Qy 721 AATCTGGCCGTTGCCAACAGGGAATTTAGTGTTCAGGGTTTCACTAAGATCAATGCC 780
|||
Db 769 AATCTGGCCGTTGCCAACAGGGAATTTAGTGTTCAGGGTTTCACTAAGATCAATGCC 828
|||
Qy 781 TTCAACTGGGCCAAGGTGGGAAAGCTGAGCTTCAAGAGAAAGGCTTCTCAATCAAGCTC 840
|||
Db 829 TTCAACTGGGCCAAGGTGGGAAAGCTGAGCTTCAAGAGAAAGGCTTCTCAATCAAGCTC 888
|||
Qy 841 CGGCGCAGATGCAAAATGATGCTACAGAGATACCTTGAATTCCTGATGGCAGTCGGGAT 900
|||
Db 889 CGGCGCAGATGCAAAATGATGCTACAGAGATACCTTGAATTCCTGATGGCAGTCGGGAT 948
|||
Qy 901 TTCTGCAAGTCTTCTGAAAAATCTGTGTGAACATCATGCTTCTTGAACCTTTTGA 960
|||
Db 949 TTCTGCAAGTCTTCTGAAAAATCTGTGTGAACATCATGCTTCTTGAACCTTTTGA 1008
|||
Qy 961 GAGCCCAAAACCAAGGCCCAAGCCGCTCTTAAAGCGGGGGTCAATTTGGGTTCACT 1020
|||
Db 1009 GAGCCCAAAACCAAGGCCCAAGCCGCTCTTAAAGCGGGGGTCAATTTGGGTTCACT 1068
|||
Qy 1021 GGTTCGGAATCAGAAAGGCTTCTGACTATGTTAAAGAGAGACATAAGAGGTGCAG 1080
|||
Db 1069 GGTTCGGAATCAGAAAGGCTTCTGACTATGTTAAAGAGAGACATAAGAGGTGCAG 1128
|||
Qy 1081 TTTGAAAGAGACACAGCAAGATTCATTATCCGAGCCTTGTCAAGCTTACAGAA 1140
|||
Db 1129 TTTGAAAGAGACACAGCAAGATTCATTATCCGAGCCTTGTCAAGCTTACAGAA 1188
|||
Qy 1141 CTGAATTCGGAAGTGTGAGAGCTCTCAGAGAGACAGAGCTTACATTTGGAGAGGT 1200
|||
Db 1189 CTGAATTCGGAAGTGTGAGAGCTCTCAGAGAGACAGAGCTTACATTTGGAGAGGT 1248
|||
Qy 1201 GCGGAATCTCAGAGGGGCGAGAGCTGCGCGAGAGAGAAAGAAACGGAAGGTTTCCGCGGG 1260
|||
Db 1249 GCGGAATCTCAGAGGGGCGAGAGCTGCGCGAGAGAGAAAGAAACGGAAGGTTTCCGCGGG 1308
|||
Qy 1261 GAGCCGGGGTCCGAGCCCTGCGCGAGAGAGAAAGCCCGCGGGTTAAACAAGCAGCG 1320
|||

|||||
Db 1309 GAGCCGGGGTCCGAGCCCGAGCTTGCGCGAGAGAAAGCCCGCGGGTTAAACAAGCAGCG 1368
|||
Qy 1321 GAGCGAGCGAGCTTGGCGGCCACGAGAGAAAGAGAGAGGTGTTAAGATAGAGCCAG 1380
|||
Db 1369 GAGCGAGCGAGCTTGGCGGCCACGAGAGAAAGAGAGAGGTGTTAAGATAGAGCCAG 1428
|||
Qy 1381 CAGAGTAAACCTCAGAGCCCGGAGCCAGACAGAGCTCTTGATCTGCAAGTCTCACTT 1440
|||
Db 1429 CAGAGTAAACCTCAGAGCCCGGAGCCAGACAGAGCTCTTGATCTGCAAGTCTCACTT 1488
|||
Qy 1441 TCCGAGCTGTCTGTGAATCTCGAGGGGGGAGTGGCCCTTGCCAAAGTGAATCTTCCC 1500
|||
Db 1489 TCCGAGCTGTCTGTGAATCTCGAGGGGGGAGTGGCCCTTGCCAAAGTGAATCTTCCC 1548
|||
Qy 1501 AACCTGAGCCCGGACACCAAGAGAGCTCTCCCTTGATCAGCCGCTGTGAATGACAG 1560
|||
Db 1549 AACCTGAGCCCGGACACCAAGAGAGCTCTCCCTTGATCAGCCGCTGTGAATGACAG 1608
|||
Qy 1561 GCTTCCCCCGGACGAGATGAGAGTGAAGGCGGAGAGAGATTCCCACTGAATAA 1620
|||
Db 1609 GCTTCCCCCGGACGAGATGAGAGTGAAGGCGGAGAGAGATTCCCACTGAATAA 1668
|||
Qy 1621 GCGTACTTCAATGCTAAGGAATGCTTACACCGAGCCGACATATCTGAAGATCTCGAA 1680
|||
Db 1669 GCGTACTTCAATGCTAAGGAATGCTTACACCGAGCCGACATATCTGAAGATCTCGAA 1728
|||
Qy 1681 GTTATCATCTGCTGTTTCAAGACACAGTGAAGAAAGAGAGCGCATGCCGGAACGACTG 1740
|||
Db 1729 GTTATCATCTGCTGTTTCAAGACACAGTGAAGAAAGAGAGCGCATGCCGGAACGACTG 1788
|||
Qy 1741 AAAAGTCTCATATTCGGAATTTTGAACCTTTGCAAAATTTCACTAATTTTCTCAG 1800
|||
Db 1789 AAAAGTCTCATATTCGGAATTTTGAACCTTTGCAAAATTTCACTAATTTTCTCAG 1848
|||
Qy 1801 GAAATTTAGCAACGACTTGCCCTGTGGGAAGGCGCTCAAAAGCCCAATCAAGATTAC 1860
|||
Db 1849 GAAATTTAGCAACGACTTGCCCTGTGGGAAGGCGCTCAAAAGCCCAATCAAGATTAC 1908
|||
Qy 1861 CAAGAATTCGCGAGTGTATGCTGAAGAACTTCAAGGCAATGAACACCTGCGGCTCAC 1920
|||
Db 1909 CAAGAATTCGCGAGTGTATGCTGAAGAACTTCAAGGCAATGAACACCTGCGGCTCAC 1968
|||
Qy 1921 CTGTGAAAGCACAGGAGGCTTGGAGGCTCTGAGAAATGGAATCAAGCTCCGCGCG 1980
|||
Db 1969 CTGTGAAAGCACAGGAGGCTTGGAGGCTCTGAGAAATGGAATCAAGCTCCGCGCG 2028
|||
Qy 1981 CTGAGGAATCTTGTGAAGACCTTGAAGCTGAAGAGTGTATTAATCTACCGCTCAACACC 2040
|||
Db 2029 CTGAGGAATCTTGTGAAGACCTTGAAGCTGAAGAGTGTATTAATCTACCGCTCAACACC 2088
|||
Qy 2041 TTCTCTGCGGCGCACTGCAACGCGCTCATGCACTAACAGAGTCTTGAGCGGCTGTGC 2100
|||
Db 2089 TTCTCTGCGGCGCACTGCAACGCGCTCATGCACTAACAGAGTCTTGAGCGGCTGTGC 2148
|||
Qy 2101 AAACACACCCGCGAGGCAAGCGGACTTCAAGGACTGCGGAGCCGCTTGGCAGAGATC 2160
|||
Db 2149 AAACACACCCGCGAGGCAAGCGGACTTCAAGGAGCTCCGAGCCGCTTGGCAGAGATC 2208
|||
Qy 2161 ACGAGATGTGGGCAACAGTCCAGGTAAGATCAAGATGGAAGAAATTTCCAGAAAGCTG 2220
|||
Db 2209 ACGAGATGTGGGCAACAGTCCAGGTAAGATCAAGATGGAAGAAATTTCCAGAAAGCTG 2268
|||
Qy 2221 CACGAATCAAGAAAGATTTGATTTGGACTTGAACAATCTTGTGTTCGCGAGAGGAGTTC 2280
|||
Db 2269 CACGAATCAAGAAAGATTTGATTTGGACTTGAACAATCTTGTGTTCGCGAGAGGAGTTC 2328
|||
Qy 2281 ATCGTCTGGGCAAGCTTCAAGAGCTTCTGGGGAAGGGCTTCAAGACGCAATGTTCTTC 2340
|||
Db 2329 ATCGTCTGGGCAAGCTTCAAGAGCTTCTGGGGAAGGGCTTCAAGACGCAATGTTCTTC 2388
|||
Qy 2341 CTGTTCAACGAGCTCTGCTATACAGAGCCGGGGGCTGACGCGCTCCAAATCAAGTTTAA 2400
|||
Db 2389 CTGTTCAACGAGCTCTGCTATACAGAGCCGGGGGCTGACGCGCTCCAAATCAAGTTTAA 2448
|||
```


QY 2401 GTCCAGGAGGAGTCCCGCTCTATGAGATGACGATTGAGAGAGCGAAGACGAGTGGGG 2460
DB 2449 GTCCAGGAGGAGTCCCGCTCTATGAGATGACGATTGAGAGAGCGAAGACGAGTGGGG 2508
QY 2461 GTGCCCCACTGCTGACCTCCGCGGCGACGCGAGTGCATCATCTGCGCGCAGTTCT 2520
DB 2509 GTGCCCCACTGCTGACCTCCGCGGCGACGCGAGTGCATCATCTGCGCGCAGTTCT 2568
QY 2521 CGGTCCGAGATGAGAGAGTGGTTGAGGACATCCAGATGCGCATTTGACTGGCGAGAG 2580
DB 2569 CGGTCCGAGATGAGAGAGTGGTTGAGGACATCCAGATGCGCATTTGACTGGCGAGAG 2628
QY 2581 AGCAGAGAGCCCGCCCTGAGTTCCTGCGCAGAGAGCCCGCTGACAAAGTCCCTGAT 2640
DB 2629 AGCAGAGAGCCCGCCCTGAGTTCCTGCGCAGAGAGCCCGCTGACAAAGTCCCTGAT 2688
QY 2641 GAGCCACGCGCGCTGACAGAGATGACGAGTGCATCTGACGCGCTGCGCAGATGCTG 2700
DB 2689 GAGCCACGCGCGCTGACAGAGATGACGAGTGCATCTGACGCGCTGCGCAGATGCTG 2748
QY 2701 GAGCCGACGCGCGCGCAGCGCGGACACATGCTGCTGCGCAGCGGACAC 2760
DB 2749 GAGCCGACGCGCGCGCAGCGCGGACACATGCTGCTGCGCAGCGGACAC 2808
QY 2761 AGCGTCCATGAGTGAATTCAGATGCGAGTGAAGATGATGTTGTTGAAACCTGCTG 2820
DB 2809 AGCGTCCATGAGTGAATTCAGATGCGAGTGAAGATGATGTTGTTGAAACCTGCTG 2868
QY 2821 AGGAAATTCAAAACAGCAACGCGGTGCGAGAGCTGTGGTGTGTTCAAACTTCTG 2880
DB 2869 AGGAAATTCAAAACAGCAACGCGGTGCGAGAGCTGTGGTGTGTTCAAACTTCTG 2928
QY 2881 CTGTTTCTTCAAAATCAACACGAGATCATCTCCCTGCGCTGCTGCTGCTGCTG 2940
DB 2929 CTGTTTCTTCAAAATCAACACGAGATCATCTCCCTGCGCTGCTGCTGCTGCTG 2988
QY 2941 TACTGCTCAACATCCCTCTGAGTCCGAGAGATCATCAAGAACTGATGCTGCTGCTG 3000
DB 2989 TACTGCTCAACATCCCTCTGAGTCCGAGAGATCATCAAGAACTGATGCTGCTGCTG 3048
QY 3001 CACTTCAAGTCCGACGCTGTAATCTTCAAGGCGGAAAGCGAGTCAAGTTCGAAAGTGG 3060
DB 3049 CACTTCAAGTCCGACGCTGTAATCTTCAAGGCGGAAAGCGAGTCAAGTTCGAAAGTGG 3108
QY 3061 ATGGAGATGATCCGAGTGCACCAAGCTCTGCTCGGAGCCCAAGTTCGAGCCACAA 3120
DB 3109 ATGGAGATGATCCGAGTGCACCAAGCTCTGCTCGGAGCCCAAGTTCGAGCCACAA 3168
QY 3121 GAGTCTCTTGTGTAT 3135
DB 3169 GAGTCTCTTGTGTAT 3183

RESULT 2
AB008430 3442 bp mRNA linear PRI 13-FEB-1999
LOCUS AB008430 Homo sapiens mRNA for CDEP, complete cds.
DEFINITION AB008430
ACCESSION AB008430.1 GI:2766164
VERSION
KEYWORDS CDEP.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens (human)

REFERENCE
AUTHORS Koyano,Y., Kawamoto,T., Shen,M., Yan,W., Noshiro,M., Fujii,K. and
Kato,Y.
TITLE Molecular cloning and characterization of CDEP, a novel human
protein containing the ezrin-like domain of the band 4.1
superfamily and the Db1 homology domain of Rho guanine nucleotide
exchange factors

JOURNAL Biochem. Biophys. Res. Commun. 241 (2), 369-375 (1997)
PUBMED 9425278
REFERENCE 2 (base 1 to 3442)
AUTHORS Koyano,Y., Kawamoto,T. and Kato,Y.
TITLE Direct Submision
JOURNAL Submitted (22-OCT-1997) Takeshi Kawamoto, Hiroshima University
School of Dentistry, Department of Biochemistry, 1-2-3 Kasumi
Minami-Ku, Hiroshima, Hiroshima 734, Japan
(E-mail:tkawamo@ipc.hiroshima-u.ac.jp, Tel:082-257-5688,
Fax:082-257-5629)

FEATURES
source
Location/Qualifiers
1..3442
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/cell_type="chondrocyte"
/tissue_type="cartilage"
/dev_stage="embryo"
49..3186
/function="Rho Guanine Nucleotide Exchange Factor"

CDS

/note="Band 4.1 superfamily"
/product="CDEP"
/protein_id="BAA24267.1"
/db_xref="GI:2766165"
/translation="MGEIRORPTPGSRGLGAPENSGISTLERGQKPPPTSGKLVSIKI
QMLDIDQAFVPPQAPRGVLDVAVNHNANLVBGDFGFEFDPHKITITWDLAKRIV
KOIRKPHVAVFVAFPPDHTQLEBILRLFLAQVQDLAQGLTCNDLSALLI
SHIVSEIDFDBALDREHLAKNKYIPQDADLEDKVEPHNHIQOTPAESPQLEI
ARLEWYGIHLRPADKREGTKINLAVANGIIVFOGPTINAFNMAKVKLSFKRFR
LILKRPDANSADVDLEFLMARSDPFCKSPWKI CVEHAFPLFEERPKPKVPLFRG
SSPRPSGRQKQVLDVYKEGHHKKVQPEERKIKTHIRLSAQPEINSEVLEOSOS
TSITTBGABESRGSGSCRGKPKYSAGSPGSHPSAPRPSRANKQADGASAPRE
ESBVADKRTQSKPQPPQSTSLTSSPHLSVSVSGVAVAPANTYLDLSPVDTKQ
ASPLISPLINDQACPTDDEDEGRRRRFTDAAYFAKSVSTTERVYLDLEVTISWF
OSTVSKEDMPALAKSLIPNPEFLPKFTNFKELIEQSLAMWEGSNQIIRYQRIQ
DVLKNIOGMKLAHLMTGSEALBELENGISRSRLNENFCRDLQKCYPLNTEFL
LPIRLHMYKQVLEIKGHPSPHADPDCRAALAEITEMVAOLHGMITKMNPKOL
HELKDLIGIDNLVYVPREFTIRLSLSLSKSGGLQOORMFPLPNDVLLYTSRGLTASNO
FVHGQPLVIGMTIESEDEKGVPHCLITLGRQSTITVAASRSEBEKVEDIQMAD
LAEKSSPAPEFLASPPDNKSPDENATADQSEBDLSASRTSLERQAFKGNATWAV
CHNRNYSVMVDFIASVENQLSGNLRFKRNKNGQKLVVVTNFCLPYKSHQDNHP
LASPLPLGSLTIPSESENIQKDYVFKLHFKSHVYVYFAESEYTERFMMEVIRSATSS
ASRPVHLSHKESLVY"

polyA_site

3442
/note="50 a nucleotides"

ORIGIN

Query Match 100.0%; Score 3135; DB 5; Length 3442;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGAGAAATAGACAGAGCGGACCCGACGATCAAGACTGGGGGCCCCGAGAAATTCG 60
DB 49 ATGGGAGAAATAGACAGAGCGGACCCGACGATCAAGACTGGGGGCCCCGAGAAATTCG 108
QY 61 GGGATCAGTACTCTTGGAAAGTGGACGAAGCCGCCCAACACCTTCAGAGAAATCTCGTG 120
DB 109 GGGATCAGTACTCTTGGAAAGTGGACGAAGCCGCCCAACACCTTCAGAGAAATCTCGTG 168
QY 121 TCCATCAAAATCCAGATGCTGATGACACCCAGAGAGCATTTGAAGTTCCACAAAGAGCT 180
DB 169 TCCATCAAAATCCAGATGCTGATGACACCCAGAGAGCATTTGAAGTTCCACAAAGAGCT 228
QY 181 CCTGGAGAGTGTGCTGATGACAGTTTCAACCACTCAACTCTGAGAGGTGACTAT 240
DB 229 CCGGAGAGTGTGCTGATGACAGTTTCAACCACTCAACTCTGAGAGGTGACTAT 288
QY 241 TTGGGCTCGAGTTTCTTATCACAAAAGATCAGGTGTGCTGATCTCTCTAAACCC 300
DB 289 TTGGGCTCGAGTTTCTTATCACAAAAGATCAGGTGTGCTGATCTCTCTAAACCC 348
QY 301 ATTGTAAACAGATTGAAGCCAAAGCAGTTGTGTTAAGTTGTGTAATTCCTTT 360

Db 1349 ATTGGAACAGATTGGAAGCCAAAGCAAGTTGTTGTTAAGTTGTGTGAATCTTT 408
Qy 361 CCGCTGACCAACACACAACTCCAGAGAAATCAACAAGTACCTGTCCGCTGCAAGTG 420
Db 409 CCGCTGACCAACACAACTCCAGAGAAATCAACAAGTACCTGTCCGCTGCAAGTG 468
Qy 421 AAGCAGACTTGGCTCAAGGAGGTGAGCTGTATATGACACAGCGACGCTCTTGATT 480
Db 469 AAGCAGACTTGGCTCAAGGAGGTGAGCTGTATATGACACAGCGACGCTCTTGATT 528
Qy 481 TCACACATTTGCAATCTGAGATTGGGGAATTTGATGAAGCTTGGACAGAGACCTTA 540
Db 529 TCACACATTTGCAATCTGAGATTGGGGAATTTGATGAAGCTTGGACAGAGACCTTA 588
Qy 541 GGAAGAAATTAATTAACCTAGCAAGACGCACTAGAGACAAATGCTGAATTTTAC 600
Db 589 GGAAGAAATTAATTAACCTAGCAAGACGCACTAGAGACAAATGCTGAATTTTAC 648
Qy 601 CATTAACCACTTGGACAAACACAGCAGATCAGATTTTCAGCTCCTAGAGATTGCCGT 660
Db 649 CATTAACCACTTGGACAAACACAGCAGATCAGATTTTCAGCTCCTAGAGATTGCCGT 708
Qy 661 CCGCTAGAGATGTATGAAATCCGGTTGCAACCGGCCAAGACAGAGGAGCAAGAAATC 720
Db 709 CCGCTAGAGATGTATGAAATCCGGTTGCAACCGGCCAAGACAGAGGAGCAAGAAATC 768
Qy 721 AATCTGCGCTTGGCCAAACCGGAAATCTAAGTCTTCAAGGTTTCACTAAGTCAATGC 780
Db 769 AATCTGCGCTTGGCCAAACCGGAAATCTAAGTCTTCAAGGTTTCACTAAGTCAATGC 828
Qy 781 TTCAACTGGGCCAAGGTGCGGAACTGAGCTTCAAGAGAGCGCTTCTCATCAAGCTC 840
Db 829 TTCAACTGGGCCAAGGTGCGGAACTGAGCTTCAAGAGAGCGCTTCTCATCAAGCTC 888
Qy 841 CCGCCAGATGCCAATATGTCGTAACAGATACCTTGAATTCCTGATGCGAGTGGGAT 900
Db 889 CCGCCAGATGCCAATATGTCGTAACAGATACCTTGAATTCCTGATGCGAGTGGGAT 948
Qy 901 TTCTGCAATCTCTTGGAAATCTGTGTGAACATCATAGCCCTTTTGAACCTTTTGA 960
Db 949 TTCTGCAATCTCTTGGAAATCTGTGTGAACATCATAGCCCTTTTGAACCTTTTGA 1008
Qy 961 GAGCCCAACCAAAAGCCCAAGCCGTCCTTTAGCCGGGGTCATCATTTCCGTTCACT 1020
Db 1009 GAGCCCAACCAAAAGCCCAAGCCGTCCTTTAGCCGGGGTCATCATTTCCGTTCACT 1068
Qy 1021 GGTCCGACTCAGAGAGGTTCTCGACTATGTTAAAGAGAGACATAAGAGGTGAC 1080
Db 1069 GGTCCGACTCAGAGAGGTTCTCGACTATGTTAAAGAGAGACATAAGAGGTGAC 1128
Qy 1081 TTGGAAGAGACAGAGAGATTCATTTATCCGAGCCTTGCCTTCAAGCCTACAGAA 1140
Db 1129 TTGGAAGAGACAGAGAGATTCATTTATCCGAGCCTTGCCTTCAAGCCTACAGAA 1188
Qy 1141 CTGAATTCGGAAGTCTGAGAGCTCTCAGCAGAGACCAAGCCTTACATTTGGAGAGGT 1200
Db 1189 CTGAATTCGGAAGTCTGAGAGCTCTCAGCAGAGACCAAGCCTTACATTTGGAGAGGT 1248
Qy 1201 GCCGAATCTCAGAGGGGCGAGAGCTCCGCGAGAGAAAGAACCGAAGTTTCCGCGGG 1260
Db 1249 GCCGAATCTCAGAGGGGCGAGAGCTCCGCGAGAGAAAGAACCGAAGTTTCCGCGGG 1308
Qy 1261 GAGCCGGGGTCCGACCCGAGCCCTGCGCGAGAGAAAGCCCGCGGGTTAAACAAGCGGG 1320
Db 1309 GAGCCGGGGTCCGACCCGAGCCCTGCGCGAGAGAAAGCCCGCGGGTTAAACAAGCGGG 1368
Qy 1321 GACGAGACCGCTCCGCGCCCAACGAGAGAGAGAGAGAGTCTTAAAGATAGAACCCAG 1380
Db 1369 GACGAGACCGCTCCGCGCCCAACGAGAGAGAGAGAGAGTCTTAAAGATAGAACCCAG 1428
Qy 1381 CAGAGTAACTTCAGCCCGCAGCCAGACAGGCTCCCTGACTGCGAGTCTCTCACTTT 1440

Db 1429 CAGAGTAACTTCAGCCCGCAGCCAGACAGAGCTCCCTGACTGCGAGTCTCACTT 1488
Qy 1441 TCCGAGCTGTCTGTAACTCCGAGGGGGAGTGGCCCTGCCCAAGCTGACTTGTCTCC 1500
Db 1489 TCCGAGCTGTCTGTAACTCCGAGGGGGAGTGGCCCTGCCCAAGCTGACTTGTCTCC 1548
Qy 1501 AACCTGAGCCCGCAGACCAAGAGGCTCTCCCTTGAATCAGCCGCTGTGAATGACAG 1560
Db 1549 AACCTGAGCCCGCAGACCAAGAGGCTCTCCCTTGAATCAGCCGCTGTGAATGACAG 1608
Qy 1561 GCTTCCCCCGGACCGAGAGATAGAGGCCGAGAGAGAGATTCCTCACTGATTA 1620
Db 1609 GCTTCCCCCGGACCGAGAGATAGAGGCCGAGAGAGAGATTCCTCACTGATTA 1668
Qy 1621 GGTACTTCACTAGCTAAGAGATGTCTACCAACGAGCAGCAATATCTGAAGATCTGAA 1680
Db 1669 GGTACTTCACTAGCTAAGAGATGTCTACCAACGAGCAGCAATATCTGAAGATCTGAA 1728
Qy 1681 GTTATCACTTCGTGTTTCAAGACAGTGAACAAAGAGACGCCATGCGGAAGCATG 1740
Db 1729 GTTATCACTTCGTGTTTCAAGACAGTGAACAAAGAGACGCCATGCGGAAGCATG 1788
Qy 1741 AAAAGTCTCATTTCCCGAATTTTGAACCTTTGCACAAAATTTCACTAATTTTCTCAAG 1800
Db 1789 AAAAGTCTCATTTCCCGAATTTTGAACCTTTGCACAAAATTTCACTAATTTTCTCAAG 1848
Qy 1801 GAAATTTGAGCAAGCACTTGCCCTGTGGGAGGCGCTCAATAGCCCAATCAGAGATTAC 1860
Db 1849 GAAATTTGAGCAAGCACTTGCCCTGTGGGAGGCGCTCAATAGCCCAATCAGAGATTAC 1908
Qy 1861 CAAAGATTCGGGATGTCAATGCTGAAGAACATTCAGGGCATAGACACTGCGGCTCAC 1920
Db 1909 CAAAGATTCGGGATGTCAATGCTGAAGAACATTCAGGGCATAGACACTGCGGCTCAC 1968
Qy 1921 CTGTGGAAGACAGAGAGGCTTGAAGGCCCTGAGAGAAATGAAATCAAGGCTCCGCGCG 1980
Db 1969 CTGTGGAAGACAGAGAGGCTTGAAGGCCCTGAGAGAAATGAAATCAAGGCTCCGCGCG 2028
Qy 1981 CTGAGAACTTCTGACAGACTTTGAGCTGACAGAGTGTGTTAATCTACCGCTCAACACC 2040
Db 2029 CTGAGAACTTCTGACAGACTTTGAGCTGACAGAGTGTGTTAATCTACCGCTCAACACC 2088
Qy 2041 TTCTCTCTGCGGCACTGACCGGCTCATGCAACAAGAGTCTCTGAGCGGCTGTGC 2100
Db 2089 TTCTCTCTGCGGCACTGACCGGCTCATGCAACAAGAGTCTCTGAGCGGCTGTGC 2148
Qy 2101 AAAACACACCCCGCAGACGACGCTTCAAGGAGCTGCGGAGCGCTTTGGCAAGATC 2160
Db 2149 AAAACACACCCCGCAGACGACGCTTCAAGGAGCTGCGGAGCGCTTTGGCAAGATC 2208
Qy 2161 ACGAGATAGTGGCAAGCTCCAGCGTACAGATGATCAAGATGAGAAATTTCCAGAAAGT 2220
Db 2209 ACGAGATAGTGGCAAGCTCCAGCGTACAGATGATCAAGATGAGAAATTTCCAGAAAGT 2268
Qy 2221 CACGAATCTCAGAAAGATTTGATGGAATCTTGTGTTCCGGGAAGGAGTTTC 2280
Db 2269 CACGAATCTCAGAAAGATTTGATGGAATCTTGTGTTCCGGGAAGGAGTTTC 2328
Qy 2281 ATCCGTTGGGAGGCTCAGCAAGCTCTCGGGAAAGGGCTCACAAGCGCATGTTCTTC 2340
Db 2329 ATCCGTTGGGAGGCTCAGCAAGCTCTCGGGAAAGGGCTCACAAGCGCATGTTCTTC 2388
Qy 2341 CTGTTCAACGAGCTCTGCTATACAGAGCCGGGGCTGACGGCTCAATCAGTTTAA 2400
Db 2389 CTGTTCAACGAGCTCTGCTATACAGAGCCGGGGCTGACGGCTCAATCAGTTTAA 2448
Qy 2401 GTTCAAGGAGCTCTCCGCTCATGAGCATGATGAGAGAGCGAAGACGAGTGGGG 2460
Db 2449 GTTCAAGGAGCTCTCCGCTCATGAGCATGATGAGAGAGCGAAGACGAGTGGGG 2508
Qy 2461 GTTCCCACTGCTGACCTCCGAGGCGAGCGGAGTCCATCATGTTGCGCCGAGTTCT 2520
Db 2509 GTTCCCACTGCTGACCTCCGAGGCGAGCGGAGTCCATCATGTTGCGCCGAGTTCT 2568

```

QY 2521 CGGTCCGAGATGAGAAAGGGGTTGAGAGATCCAGATGGCCATTGACCTGGCGGAGAAAG 2580
DB 2569 CGGTCCGAGATGAGAAAGGGGTTGAGAGATCCAGATGGCCATTGACCTGGCGGAGAAAG 2628
QY 2581 AGCAGACGCCCCCCTGAGTTCCTGGCCAGACGCCCTCTGACAAACAAGTCCCTGAT 2640
DB 2629 AGCAGACGCCCCCCTGAGTTCCTGGCCAGACGCCCTCTGACAAACAAGTCCCTGAT 2688
QY 2641 GAAGCCACCGCGGCTGACCAAGATGACCTGAGCCCTTCGCGACATGCTG 2700
DB 2689 GAAGCCACCGCGGCTGACCAAGATGACCTGAGCCCTTCGCGACATGCTG 2748
QY 2701 GAGCGCCAGGCCCCCGACCGCGGCAACAAATGATGAGTGGCTGGGACCGGCAACACC 2760
DB 2749 GAGCGCCAGGCCCCCGACCGCGGCAACAAATGATGAGTGGCTGGGACCGGCAACACC 2808
QY 2761 AGCGTCTCATGAGTGAATTCAGCATCGCAGTGGAGATGATGATGTTGTTGAAACCTGCTG 2820
DB 2809 AGCGTCTCATGAGTGAATTCAGCATCGCAGTGGAGATGATGATGTTGTTGAAACCTGCTG 2868
QY 2821 AGGAAATTCAAAAACGCAACGGGTGGCAGAACCTGTGGTGTTCACAAACTTCTGC 2880
DB 2869 AGGAAATTCAAAAACGCAACGGGTGGCAGAACCTGTGGTGTTCACAAACTTCTGC 2928
QY 2881 CTGTTCTTCTACAAATCAGACACGAGCAATCATCCCTTGGCCAGCTGCTGCTCGC 2940
DB 2929 CTGTTCTTCTACAAATCAGACACGAGCAATCATCCCTTGGCCAGCTGCTGCTCGC 2988
QY 2941 TACTCGTCACCATCCCTCTGTGAGTCCGAGAACATCCAGAAAGCTACGTTCAAGCTG 3000
DB 2989 TACTCGTCACCATCCCTCTGTGAGTCCGAGAACATCCAGAAAGCTACGTTCAAGCTG 3048
QY 3001 CACTTCAAGTCCAGCTCTACTACTTCAAGGCGGAAAACGAGTACACGTTGAAAGGTG 3060
DB 3049 CACTTCAAGTCCAGCTCTACTACTTCAAGGCGGAAAACGAGTACACGTTGAAAGGTG 3108
QY 3061 ATGGAAGTATCCGCGAGTGCACACACTCTGCTCCGACCCCAAGCTTTGAGCCACAA 3120
DB 3109 ATGGAAGTATCCGCGAGTGCACACACTCTGCTCCGACCCCAAGCTTTGAGCCACAA 3168
QY 3121 GAGTCTCTGTGTAT 3135
DB 3169 GAGTCTCTGTGTAT 3183

```

RESULT 3
BC041595 3546 bp mRNA linear PRI 08-MAR-2005
LOCUS Homo sapiens FERM, Rhogef (ARHGEF) and plectectrin domain protein 1
DEFINITION (Chondrocyte-derived), transcript variant 1, mRNA (cDNA clone MGC:52223 IMAGE:5443230), complete cds.
ACCESSION BC041595
VERSION BC041595.1 GI:27370806
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens (human)
REFERENCE 1 (bases 1 to 3546)
AUTHORS Strunberg RL, Feingold EA, Grouse LH, Derge JG, Klausner RD, Collins FS, Wagner L, Shennan CM, Schuler GD, Altshul SF, Zeeberg B, Buetow KH, Schaefer CF, Bhat NK, Hopkins RF, Jordan H, Moore T, Max SI, Wang J, Hsieh F, Diatchenko L, Marusina K, Farmer AA, Rubin GM, Hong L, Stapleton M, Soares MB, Bonaldo MF, Casavant TL, Scheetz TE, Brownstein WJ, Ueda N, Toshiyuki S, Carninci P, Prange C, Raha SS, Loquellano NA, Peters GJ, Abramson RD, Mullaby SJ, Bosak SA, McGowan PJ, McKernan KJ, Malek JA, Gunaratne PH, Richardson S, Worley KC, Hale S, Garcia AM, Gay LJ, Holik SM, Villalón DK, Wuzny DM, Sodergren EJ, Lu X, Gibbs RA, Fahy J, Helton E, Kesteman M, Madan A, Rodrigues S, Sanchez A, Whitting M, Madan A, Young AC, Shevchenko Y, Bouffard GG, Blakeley RW,

Touchman JW, Green ED, Dickson MC, Rodriguez AC, Grimwood J, Schmutz J, Myers RM, Butterfield JS, Krzywinski MI, Skelske U, Smalius DE, Schmech A, Schein JE, Jones SJ and Marra MA. Mammalian Gene Collection Program Team
 Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 1247932
 2 (bases 1 to 3546)
REFERENCE Director MGC Project.
AUTHORS Direct Submission
TITLE Submitted (20-DEC-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
COMMENT NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgabs-rc@mail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC),
 Gaithersburg, Maryland;
 Web site: <http://www.nisc.nih.gov/>
 Contact: nisc.mgc@nih.gov
 Akhter, N., Ayle, K., Beckstrom-Stenberg, S.M., Benjamin, B., Blakeley, R.W., Bouffard, G.G., Breen K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Lart, P., Legaspi, R., Maduro, O.L., Masello, C., Maskeri, B., Mastrian, S.D., McLokey, J.C., McDowell, J., Pearson, R., Stantrop, S., Thomas, P.J., Touchman, J.W., Tsugeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
 Series: IRRL Plate: 44 Row: 1 Column: 24.
FEATURES
source
 1..3546
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:52223 IMAGE:5443230"
 /issue_type="Brain, astrocytoma, grade IV"
 /clone_id="N1H_MGC_98"
 /lab_host="DH10B"
 /note="Vector: pOTB7"
 1..3546
 /gene="PARP1"
 /note="BYNAME: CDEP, MGC87400"
 /db_xref="GeneID:10160"
 /db_xref="MIM:602654"
 266..3403
 /gene="PARP1"
 /codon_start=1
 /product="FERM, Rhogef, and plectectrin domain protein 1, isoform 1"
 /protein_id="AAH41595.1"
 /db_xref="GI:49898405"
 /db_xref="GeneID:10160"
 /db_xref="MIM:602654"
 /translation="MGEIRQRTPSGRGLAPENSGLSTERGOKPEPTPSGKLVSTIKI QMLDTPQAFVQAPQPGVLDVAVCNHNLVBDYFGJLFPDHKITYMLDLKRTY KQIRPKVHVAVKPEPPDTOLQEBLRYLFLAQVQDLAOGKLTNDVSALLI SHIVSEIDPDEALDRBHLAKNKYIPQDALLDRIYFHHNHIGQTPAESFOLLEI ARLEWYGRILPADEBETKINLAVANTGILVFOGFTKINAFNAKVKLSFKRKF LKRLPDANSAVODTLEPLMASRDECKSPFKICVHNHAFPEBPKPKPKVLSRGS SFRPSGRQKQVLDYKVGCGKTKQVPRKHSIKISIRLSAOPTEINSEVLEQSOO TSITRGEASPEGQSCRRGKPKYSAGPSGHPAPRPSRPAKNGKADGAASATPEE BEEVKDRQOSKPPQPPSTSLGSPHLSVNSQGVAPANVTLSPLSPDKO ASPLSPLANDACPTDEDEGRRRRPTDRAVYIAKSVSTTERLYLQDLVITSWF GSTVSKEDAMPALKSLIPNPEPLHKFTNLEKEIEQLALMEGRSNAQIRDYRIG

DVLMKNIQGMKHLAAHLWKHSBALBLENKISRSRLNENFCHDFELOKVCYPLNTFL
LAEHLRMHYKQVLEBLCGHPSPSHADPFDCALALBETBMAVLAHGTMIKMENTOKL
HELKDLIGIDNLYVGRREFIRLSXISGLQOEMFELNDVLLATGSRLESLTASNO
FVYHQDPLVYGMTIERSEBDEBKVPHCTLRGROSTIYVASSRSEBKEKVEDIQMAID
LAEKSSPAEPFLASPDNDKSPDEBTADDOSEEDLSRSTLSERQAPHRANTVHV
CWHRTSVSMVDFSIAVENQLSRKKNNGWOKLWVNTNLCFLFYKSHQONHP
LASLPLGYSLTIPSEBENIOKDYVFKLHFKSHVYVFRASEYTEREWMVEVRSATSS
ASRPHVLSHKESTLVY"

ORIGIN

Query Match	100.0%	Score 3135	DB 5	Length 3546
Best Local Similarity	100.0%	Pred. No. 0		
Matches 3135	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY	1	ATGGGAGAAATAGAGACAGCGCCGACCCGAGATCAACGACTGGGGGCCCCGGAAAAATTCG	60	
Db	266	ATGGGAGAAATAGAGACAGCGCCGACCCGAGATCAACGACTGGGGGCCCCGGAAAAATTCG	325	
QY	61	GGGATCAGTACCTTGGAAAGTGGACAGAAAGCCGCCCAACACCTTCAGAAAACTCGTG	120	
Db	326	GGGATCAGTACCTTGGAAAGTGGACAGAAAGCCGCCCAACACCTTCAGAAAACTCGTG	385	
QY	121	TCATATCAAAATCCAGATGCTGATGACACCCAGAGAGCATTTGAAGTTCCAAAGAGCT	180	
Db	386	TCATATCAAAATCCAGATGCTGATGACACCCAGAGAGCATTTGAAGTTCCAAAGAGAGCT	445	
QY	181	CGTGGGAAGGTGCTGATGATGAGTTTGCAACCACTCAACCTGCTGGAGGTGACTAT	240	
Db	446	CGTGGGAAGGTGCTGATGATGAGTTTGCAACCACTCAACCTGCTGGAGGTGACTAT	505	
QY	241	TTTGGCTCGAGTTTCTGATCACAAAAGATCAACGCTGTGCTGATCTCTTAAACC	300	
Db	506	TTTGGCTCGAGTTTCTGATCACAAAAGATCAACGCTGTGCTGATCTCTTAAACC	565	
QY	301	ATTGTGAAAACAGATTGAAAGGCCAAAGCAAGTTGTTGTTAAGTTGTGGTAAATCTTT	360	
Db	566	ATTGTGAAAACAGATTGAAAGGCCAAAGCAAGTTGTTGTTAAGTTGTGGTAAATCTTT	625	
QY	361	CGGCTGACACACCACTCCAGAAAGACTCAAGAGTACTGTTGCGCGCTGACAGTG	420	
Db	626	CGGCTGACACACCACTCCAGAAAGACTCAAGAGTACTGTTGCGCGCTGACAGTG	685	
QY	421	AAGCAGAGCTTGCTCAAGCAGAGTTGACGTGTATGACACCAAGCAGCTCTTTGATT	480	
Db	686	AAGCAGAGCTTGCTCAAGCAGAGTTGACGTGTATGACACCAAGCAGCTCTTTGATT	745	
QY	481	TCACACATTTGCAATCTGAGATTTGGGGATTTTGAAGAGCTTTGACAGAGACCTTA	540	
Db	746	TCACACATTTGCAATCTGAGATTTGGGGATTTTGAAGAGAGCTTTGACAGAGACCTTA	805	
QY	541	GCAAAAATAAATPACTACCTCAGCAAGAGCGCACTGAGAGACAAATGCTGGAATTTTCA	600	
Db	806	GCAAAAATAAATPACTACCTCAGCAAGAGCGCACTGAGAGACAAATGCTGGAATTTTCA	865	
QY	601	CATTAACCACTTGAACAAACACGAGCAGATCAGATTTTCAGCTCCTAGAGATTGCCGT	660	
Db	866	CATTAACCACTTGAACAAACACGAGCAGATCAGATTTTCAGCTCCTAGAGATTGCCGT	925	
QY	661	CGGCTAGAGATGTATGGAATTCGGGTGGACCCGGCCCAAGACAGAGAGCAGAAAGATC	720	
Db	926	CGGCTAGAGATGTATGGAATTCGGGTGGACCCGGCCCAAGACAGAGAGCAGAAAGATC	985	
QY	721	AATCTGGCGCTTGGCAACAGGGAATTTCTAGTGTTCAGGGTTTCTAAGATCAATGCC	780	
Db	986	AATCTGGCGCTTGGCAACAGGGAATTTCTAGTGTTCAGGGTTTCTAAGATCAATGCC	1045	
QY	781	TTTCAACTGGGCCAAGGTGCGAAGCTGAGCTTCAAGAGAAAGCGCTTTCTCACTCAAGCTC	840	
Db	1046	TTTCAACTGGGCCAAGGTGCGAAGCTGAGCTTCAAGAGAAAGCGCTTTCTCACTCAAGCTC	1105	
QY	841	CGGCGAGATGCCAATATGTCGCTTCCAGATACCTTTGGAATTTCTGATGGCAGTCCGGAT	900	
Db	1106	CGGCGAGATGCCAATATGTCGCTTCCAGATACCTTTGGAATTTCTGATGGCAGTCCGGAT	1165	

QY	901	TTCTCAAGTCCCTTCTGGAAAACTGTGTGAACATCATGCTTCTTTAGACTTTTGA	960
Db	1166	TTCTCAAGTCCCTTCTGGAAAACTGTGTGAACATCATGCTTCTTTAGACTTTTGA	1225
QY	961	GAGCCAAACCAAGCCCAAGCCGCTCTTTAGCCGGGGTCAATCATTTGGTTCACT	1020
Db	1226	GAGCCAAACCAAGCCCAAGCCGCTCTTTAGCCGGGGTCAATCATTTGGTTCACT	1285
QY	1021	GGTGGACTCGAAGACAGATTCTGACTATGTATAAGAGAGACATPAAGAGGTGACG	1080
Db	1286	GGTGGACTCGAAGACAGATTCTGACTATGTATAAGAGAGACATPAAGAGGTGACG	1345
QY	1081	TTTGAAGAGAAAGCAGACAGAAATTCATTCGAGACCTTGTCTTCAAGCTTACAGA	1140
Db	1346	TTTGAAGAGAAAGCAGACAGAAATTCATTCGAGAGCTTGTCTTCAAGCTTACAGA	1405
QY	1141	CTGAATTCGAAAGTCTGAGACAGTCTCAGACAGACCAAGCTTACATTTGAGAGGT	1200
Db	1406	CTGAATTCGAAAGTCTGAGACAGTCTCAGACAGACCAAGCTTACATTTGAGAGGT	1465
QY	1201	GCCGAATCTCCAGGGGGGCAAGCTGCGCGAGAGAAAGGAAACGAAAGTTTCCCGCGG	1260
Db	1466	GCCGAATCTCCAGGGGGGCAAGCTGCGCGAGAGAAAGGAAACGAAAGTTTCCCGCGG	1525
QY	1261	GAGCCGGGTCGACCCGAGCCCTGCGCGAGAGAAACCCCGCGGTTAACAGAGGCG	1320
Db	1526	GAGCCGGGTCGACCCGAGCCCTGCGCGAGAGAAACCCCGCGGTTAACAGAGGCG	1585
QY	1321	GACGAGCCGCTCGGCCCCCAAGAGAAAGAGAGGTCTTTAAGATTGAAACCCAG	1380
Db	1586	GACGAGCCGCTCGGCCCCCAAGAGAAAGAGAGGTCTTTAAGATTGAAACCCAG	1645
QY	1381	CAGAGTAAACCTAAGCCCCCGAGCAAGGCTCCCTGACTGGAGTCTCAACCTT	1440
Db	1646	CAGAGTAAACCTAAGCCCCCGAGCAAGGCTCCCTGACTGGAGTCTCAACCTT	1705
QY	1441	TCGAGCTGTCTGTAGACTCGCAGGGGGAGTGCGCCCTGCAACGTGACTTGTCTCC	1500
Db	1706	TCGAGCTGTCTGTAGACTCGCAGGGGGAGTGCGCCCTGCAACGTGACTTGTCTCC	1765
QY	1501	AACTGAGCCCGACACCAAGCAGGCTCTCTCTTGAATGACGCCGCTGTAATGACAG	1560
Db	1766	AACTGAGCCCGACACCAAGCAGGCTCTCTCTTGAATGACGCCGCTGTAATGACAG	1825
QY	1561	GCTGCCCCCGGAGGAGATGAGAGTGAGGGCGGAGAGAGATTCCTCAACTGATAA	1620
Db	1826	GCTGCCCCCGGAGGAGATGAGAGTGAGGGCGGAGAGAGATTCCTCAACTGATAA	1885
QY	1621	GCGTACTTCAATGACTAAGAAAGTGTCTACACCGAGCAACATATCTGAAGATCTGAA	1680
Db	1886	GCGTACTTCAATGACTAAGAAAGTGTCTACACCGAGCAACATATCTGAAGATCTGAA	1945
QY	1681	GTTATCACTTGTGTTTCAAGACAGTGAACAAAGAGACGCAATGCGGAGACATCG	1740
Db	1946	GTTATCACTTGTGTTTCAAGACAGTGAACAAAGAGACGCAATGCGGAGACATCG	2005
QY	1741	AAAAGTCAATATTCGGAATTTTGAACCTTTGCAACAAATTCATTAATTTTCTCAAG	1800
Db	2006	AAAAGTCAATATTCGGAATTTTGAACCTTTGCAACAAATTCATTAATTTTCTCAAG	2065
QY	1801	GAATTTGAGCAACGACTTGCCCTGTGGAAAGCCGCTCAATATGCCAATCAGAGATTAC	1860
Db	2066	GAATTTGAGCAACGACTTGCCCTGTGGAAAGCCGCTCAATATGCCAATCAGAGATTAC	2125
QY	1861	CAAGAAATCGGAGATGTCAATGCTGAAGAAACATTCAGGCGATGAGACACTGGCGCTCAC	1920
Db	2126	CAAGAAATCGGAGATGTCAATGCTGAAGAAACATTCAGGCGATGAGACACTGGCGCTCAC	2185
QY	1921	CTGTGGAAGCAGAGGAGGCTTGGAGGCTTGGAGAAATGGAATTAAGAGCTCCGGCGGG	1980
Db	2186	CTGTGGAAGCAGAGGAGGCTTGGAGGCTTGGAGAAATGGAATTAAGAGCTCCGGCGGG	2245

QY 1981 CTGGAGAACTTCTGACAGACTTTAGCTGCAGAGAGTGTATTACTACCTACCGCTCAACACC 2040
 DB 2246 CTGGAGAACTTCTGACAGACTTTAGCTGCAGAGAGTGTATTACTACCTACCGCTCAACACC 2305
 QY 2041 TTCTCTCTGCGGCGACCTGACACCGGCTCATGCTACCAAGAGCTCTGAGCGGCTGTGC 2100
 DB 2306 TTCTCTCTGCGGCGACCTGACACCGGCTCATGCTACCAAGAGAGCTCTGAGCGGCTGTGC 2365
 QY 2101 AAACACACACCCCGGACCCGACAGCTTTCAGGAGCTGCCGACCGCTTTGGCAGAGATC 2160
 DB 2366 AAACACACACCCCGGACCCGACAGCTTTCAGGAGCTGCCGACCGCTTTGGCAGAGATC 2425
 QY 2161 ACCGAGATGTGGACAGCTCCACGATACGATGATCAAGATGAGAAATTTCCAGAACTG 2220
 DB 2426 ACCGAGATGTGGACAGCTCCACGATACGATGATCAAGATGAGAAATTTCCAGAACTG 2485
 QY 2221 CACGAATCAAGAAAGATTTGATTGGCAATTCATCTTGTGTTCCGGGAAAGGAGTTTC 2280
 DB 2486 CACGAATCAAGAAAGATTTGATTGGCAATTCATCTTGTGTTCCGGGAAAGGAGTTTC 2545
 QY 2281 ATCCGCTGTGGGACGCTCAGCAAGCTCTCGGGGAAAGGGGCTCCAGCAGCGCATGTTCTTC 2340
 DB 2546 ATCCGCTGTGGGACGCTCAGCAAGCTCTCGGGGAAAGGGGCTCCAGCAGCGCATGTTCTTC 2605
 QY 2341 CTGTTCAACGACGCTCTGCTATACAGAGCCGGGGGCTGACGGGCTCCAAATCAAGTTTAAA 2400
 DB 2606 CTGTTCAACGACGCTCTGCTATACAGAGCCGGGGGCTGACGGGCTCCAAATCAAGTTTAAA 2665
 QY 2401 GTCCACGGGAGCTCCCGCTCTATGCGATGACGATTTGAGAGAGAGCAGAGTGGGG 2460
 DB 2666 GTCCACGGGAGCTCCCGCTCTATGCGATGACGATTTGAGAGAGAGCAGAGTGGGG 2725
 QY 2461 GTGCCCCACTGCTGACACCTTCGCGGGCCAGCGGACATCATCTGTGCGCCGCAATTTCT 2520
 DB 2726 GTGCCCCACTGCTGACACCTTCGCGGGCCAGCGGACATCATCTGTGCGCCGCAATTTCT 2785
 QY 2521 CGGTCCGAGATGAGAGATGGGTTGAGAGATCATCAGATGAGCATTTGACTGTGGCGAGAG 2580
 DB 2786 CGGTCCGAGATGAGAGATGGGTTGAGAGATCATCAGATGAGCATTTGACTGTGGCGAGAG 2845
 QY 2581 AGCAGACGCCCGCCCTGAGTTCTGTGGCAGACGCCCTCTGACAAACAGTCCCTGAT 2640
 DB 2846 AGCAGACGCCCGCCCTGAGTTCTGTGGCAGACGCCCTCTGACAAACAGTCCCTGAT 2905
 QY 2641 GAAGCCACCGCGGCTGACCAAGAGTACAGAGATGACTGAGCGCTGTGGCCACATGCTGTG 2700
 DB 2906 GAAGCCACCGCGGCTGACCAAGAGTACAGAGATGACTGAGCGCTGTGGCCACATGCTGTG 2965
 QY 2701 GAGCGCACGGCCCGGACCGGCAACAAATGATGAGTGTGTGGCCACCGCAACACC 2760
 DB 2966 GAGCGCACGGCCCGGACCGGCAACAAATGATGAGTGTGTGGCCACCGCAACACC 3025
 QY 2761 AGCGTCTCCATGTGTGACTTCAAGCATGCAAGTGAAGATGATGTTGTGAAAACCTGCTG 2820
 DB 3026 AGCGTCTCCATGTGTGACTTCAAGCATGCAAGTGAAGATGATGTTGTGAAAACCTGCTG 3085
 QY 2821 AGGAATTTCAAAAACAGCAACGGGTGGCAGAAAGTGTGGGTGTTCACAAAACCTTCTGC 2880
 DB 3086 AGGAATTTCAAAAACAGCAACGGGTGGCAGAAAGTGTGGGTGTTCACAAAACCTTCTGC 3145
 QY 2881 CTGTTCTTCTCAAAATCAACAGCAAGCAATCATCCCTTGCAGAGCTGCTGTGCTCGGC 2940
 DB 3146 CTGTTCTTCTCAAAATCAACAGCAAGCAATCATCCCTTGCAGAGCTGCTGTGCTCGGC 3205
 QY 2941 TACTGCTCAACATCCCTTCTGATGCTCGAAGAACATCCAGAAAGATCACTGTTCAGACTG 3000
 DB 3206 TACTGCTCAACATCCCTTCTGATGCTCGAAGAACATCCAGAAAGATCACTGTTCAGACTG 3265
 QY 3001 CACTTCAAGTCCACAGCTCTACTTCAAGGGCGGAAAGCGAGTACAGCTTTCAGAAAGTGG 3060
 DB 3266 CACTTCAAGTCCACAGCTCTACTTCAAGGGCGGAAAGCGAGTACAGCTTTCAGAAAGTGG 3325
 QY 3061 ATGGAATGATCCGAGTGCACACAGCTGTGCTCGGACCCCAAGTGTGAGCCACAAA 3120

DB 3326 ATGGAATGATCCGCAAGTGCACCAAGCTGTGCTCGGACCCCAAGTGTGAGCCACAAA 3385
 QY 3121 GAGTCTCTGTGTAT 3135
 DB 3386 GAGTCTCTGTGTAT 3400
 RESULT 4
 CS130723 5028 bp DNA linear PAT 02-AUG-2005
 LOCUS
 DEFINITION Sequence 9 from Patent WO2005064009.
 ACCESSION CS130723
 VERSION CS130723.1 GI:71792793
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homnidae; Homo.
 REFERENCE
 1 Ostrinofc, T.P.
 AUTHORS Classification of cancer
 TITLE Patent: WO 2005064009-A 9 14-JUL-2005;
 JOURNAL Aros Applied Biotechnology APS (DX)
 FEATURES
 source
 1..5028
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 /note="FERM, RhogEF (ARHGEF) and pleckstrin domain pro"
 ORIGIN
 Query Match 99.9%; Score 3133.4; DB 2; Length 5028;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 3134; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 ATGGGAGAAATAGAGAGAGCCGACCCAGATGACAGCTGGGGGCCCCGGGAAATTCG 60
 DB 337 ATGGGAGAAATAGAGAGAGCCGACCCAGATGACAGCTGGGGGCCCCGGGAAATTCG 396
 QY 61 GGGATCAGTACCTTGGAGAGTGGACAGAAAGCCGCCCAACACTTCAGGAAATCTGTG 120
 DB 397 GGGATCAGTACCTTGGAGAGTGGACAGAAAGCCGCCCAACACTTCAGGAAATCTGTG 456
 QY 121 TCCATCAAAATCCAGATGCTGATGACACCCAGAGAGCAATTGAAGTTCCACAAAGACT 180
 DB 457 TCCATCAAAATCCAGATGCTGATGACACCCAGAGAGCAATTGAAGTTCCACAAAGACT 516
 QY 181 CCGGGAAGTGTGCTGATGATGAGATGAGTTGCAACCACTCAACTCGTGAAGGTACTAT 240
 DB 517 CCGGGAAGTGTGCTGATGATGAGATGAGTTGCAACCACTCAACTCGTGAAGGTACTAT 576
 QY 241 TTGGGCTCGAGTTTCTGATCACAAAAAGATCAGGTGTGCTGATCTCTTAAACCC 300
 DB 577 TTGGGCTCGAGTTTCTGATCACAAAAAGATCAGGTGTGCTGATCTCTTAAACCC 636
 QY 301 ATTGTAACAAGTTTGAAGGCCCAAGCAAGTTTGTGTTAAGTTGTGTGAAATTCCTT 360
 DB 637 ATTGTAACAAGTTTGAAGGCCCAAGCAAGTTTGTGTTAAGTTGTGTGAAATTCCTT 696
 QY 361 CGGCGTGAACACACAACTCCAGAGAGAACTCAGAAAGTACTGTTCCGCGTGCAGGTTG 420
 DB 697 CGGCGTGAACACACAACTCCAGAGAGAACTCAGAAAGTACTGTTCCGCGTGCAGGTTG 756
 QY 421 AAGCAGACTTGGCTCAAGGAGGTTGACGTGTAATGACAGCAGCAGCTCTTGATT 480
 DB 757 AAGCAGACTTGGCTCAAGGAGGTTGACGTGTAATGACAGCAGCAGCTCTTGATT 816
 QY 481 TCACACATTTGTGCAATCTGAGATTTGGGATTTTGTATGAAGCCTTGGACAGAGACCTTA 540
 DB 817 TCACACATTTGTGCAATCTGAGATTTGGGATTTTGTATGAAGCCTTGGACAGAGACCTTA 876

QY 541 GCAAAAATAAATACATACCTCAGCAAGACGCACTAGAGAGCAAAATCGTGAATTTTCAC 600
| | | | |
Db 877 GCAAAAAAATAAATACCTACCTCAGCAAGACGCACTAGAGAGCAAAATCGTGAATTTTCAC 936
| | | | |
QY 601 CATTAACCAATTGGACAAACACCAAGCAATCAGATTTTCCAGTCTTAGAGATTGCCGT 660
| | | | |
Db 937 CATTAACCAATTGGACAAACACCAAGCAATCAGATTTTCCAGTCTTAGAGATTGCCGT 996
| | | | |
QY 661 CGGCTAGAGATGATGAAATCGGTTGCAACCGGCGCAAGAGACAGGGAGGCAAGAAATC 720
| | | | |
Db 997 CGGCTAGAGATGATGAAATCGGTTGCAACCGGCGCAAGAGACAGGGAGGCAAGAAATC 1056
| | | | |
QY 721 AATCTGCGGTTGCCAACAAGGAAATTTAGTGTTCAGGGTTTCACTAAGATCAATGCC 780
| | | | |
Db 1057 AATCTGCGGTTGCCAACAAGGAAATTTAGTGTTCAGGGTTTCACTAAGATCAATGCC 1116
| | | | |
QY 781 TTCAACTGGGCGCAAGGTGGGAACTGAGCTTCAAGAGAAAGCGCTTTTCATCAAGCTC 840
| | | | |
Db 1117 TTCAACTGGGCGCAAGGTGGGAACTGAGCTTCAAGAGAAAGCGCTTTTCATCAAGCTC 1176
| | | | |
QY 841 CGGCGAGATGCCAATAGTGGTACAGAGATACCTTGAATTCCTGATGGCAAGTCCGAGAT 900
| | | | |
Db 1177 CGGCGAGATGCCAATAGTGGTACAGAGATACCTTGAATTCCTGATGGCAAGTCCGAGAT 1236
| | | | |
QY 901 TTCTGCAAGTCTTCTGAAAAATCTGTGTGAACATCATGCTCTTTAGACTTTTGA 960
| | | | |
Db 1237 TTCTGCAAGTCTTCTGAAAAATCTGTGTGAACATCATGCTCTTTAGACTTTTGA 1296
| | | | |
QY 961 GAGCCCAAAACCAAGCCCAAGCCCGCTCTTTAGCCGGGGGTCATCATTTGGTTCAAT 1020
| | | | |
Db 1297 GAGCCCAAAACCAAGCCCAAGCCCGCTCTTTAGCCGGGGGTCATCATTTGGTTCAAT 1356
| | | | |
QY 1021 GGTGGACTCAGAAAGGTTCTGACTATGTAAAGAGAGACATTAAGAAAGTGCAG 1080
| | | | |
Db 1357 GGTGGACTCAGAAAGGTTCTGACTATGTAAAGAGAGACATTAAGAAAGTGCAG 1416
| | | | |
QY 1081 TTGAAAGAGACAGCAAGATTCATTATCCGAGCTTGTCTTACAGCTTACAGAA 1140
| | | | |
Db 1417 TTGAAAGAGACAGCAAGATTCATTATCCGAGCTTGTCTTACAGCTTACAGAA 1476
| | | | |
QY 1141 CTGAATTCGAAATGTGTGAGAGAGTCTCAGACAGACCAAGCTTACATTTGAGAAAGT 1200
| | | | |
Db 1477 CTGAATTCGAAATGTGTGAGAGAGTCTCAGACAGACCAAGCTTACATTTGAGAAAGT 1536
| | | | |
QY 1201 GCCGAATCTCAGAGGGGCGCAGAGCTCGCGCAGAGAAAGAAACGAAAGTTTCCGCGG 1260
| | | | |
Db 1537 GCCGAATCTCAGAGGGGCGCAGAGCTCGCGCAGAGAAAGAAACGAAAGTTTCCGCGG 1596
| | | | |
QY 1261 GAGCCGGGGTCGACCCGAGCCCTGCGCCGAGAGAAAGCCCGCGGGTAAACAAGCAGGCG 1320
| | | | |
Db 1597 GAGCCGGGGTCGACCCGAGCCCTGCGCCGAGAGAAAGCCCGCGGGTAAACAAGCAGGCG 1656
| | | | |
QY 1321 GACGAGAGCCGCTCCGGCGCCACGAGAGAAAGAGAGAGAGTGTGTAAGATAGAACCCG 1380
| | | | |
Db 1657 GACGAGAGCCGCTCCGGCGCCACGAGAGAAAGAGAGAGAGTGTGTAAGATAGAACCCG 1716
| | | | |
QY 1381 CAGAATTAACCTCAGAGCCCGCAGCAGCAAGCTCCCTGACTGGCAGTCTCTCACTT 1440
| | | | |
Db 1717 CAGAATTAACCTCAGAGCCCGCAGCAGCAAGCTCCCTGACTGGCAGTCTCTCACTT 1776
| | | | |
QY 1441 TCCGAGCTGTCTGTGAATCTCGCAGGGGGAGTGGGCTCTGCAACGTGACTTGTCTCC 1500
| | | | |
Db 1777 TCCGAGCTGTCTGTGAATCTCGCAGGGGGAGTGGGCTCTGCAACGTGACTTGTCTCC 1836
| | | | |
QY 1501 AACCTGAAGCCCGCAGACCAAGCAGGCTCTCCCTTGAATCAGCCCGTGTGTAATGACAG 1560
| | | | |
Db 1837 AACCTGAAGCCCGCAGACCAAGCAGGCTCTCCCTTGAATGACAGCCCGTGTGTAATGACAG 1896
| | | | |
QY 1561 GCTTGCCCCCGCAGCAGATGAGATGAGGGCCGAGAGAAAGATTTCCCACTGATAA 1620
| | | | |
Db 1897 GCTTGCCCCCGCAGCAGATGAGATGAGGGCCGAGAGAAAGATTTCCCACTGATAA 1956
| | | | |
QY 1621 GCGTACTTATAGCTAAGAAAGTGTCTACCAAGCAGAAATATCTGAAGATCTCGAA 1680
| | | | |

Db 1957 GCGTACTTATAGCTAAGAAAGTGTCTACCAAGCAGCAACATATCTGAAGATCTCGAA 2016
| | | | |
QY 1681 GTTATCACTTGTGGTTTCAAGACCAATGAGCAAAAGAGAGCCGATCCGGAACAACG 1740
| | | | |
Db 2017 GTTATCACTTGTGGTTTCAAGACCAATGAGCAAAAGAGAGCCGATCCGGAACAACG 2076
| | | | |
QY 1741 AAAAGTCAATATTCGCAATTTTGAACCTTTGCAAAATTTTCACTAATTTTCTCAAG 1800
| | | | |
Db 2077 AAAAGTCAATATTCGCAATTTTGAACCTTTGCAAAATTTTCACTAATTTTCTCAAG 2136
| | | | |
QY 1801 GAAATTTGAGCAACGACTTGCCTGTGGAAAGCCGCTCAATATGCCCAATCAGAGATTAC 1860
| | | | |
Db 2137 GAAATTTGAGCAACGACTTGCCTGTGGAAAGCCGCTCAATATGCCCAATCAGAGATTAC 2196
| | | | |
QY 1861 CAAGAATCGGGAATGATAGTGAAGAACTTCAGGGCAATGAGACCTGGCGGCTCAC 1920
| | | | |
Db 2197 CAAGAATCGGGAATGATAGTGAAGAACTTCAGGGCAATGAGACCTGGCGGCTCAC 2256
| | | | |
QY 1921 CTGTGGAAGCACAGCGAGGCTTGGAGGCTTGGAAATGGAATCAGAGCTCCCGCGG 1980
| | | | |
Db 2257 CTGTGGAAGCACAGCGAGGCTTGGAGGCTTGGAAATGGAATCAGAGCTCCCGCGG 2316
| | | | |
QY 1981 CTGGAATCTTCTGCAAGACTTTGAGCTGCAAGAGGTGTGTACTTACCGCTCAACACC 2040
| | | | |
Db 2317 CTGGAATCTTCTGCAAGACTTTGAGCTGCAAGAGGTGTGTACTTACCGCTCAACACC 2376
| | | | |
QY 2041 TTCTCTCGCGGCACTGACCGGCTCAGTACCTCAAGACAGGTCTGAGAGGCTGTGTC 2100
| | | | |
Db 2377 TTCTCTCTCGCGGCACTGACCGGCTCAGTACCTCAAGACAGGTCTGAGAGGCTGTGTC 2436
| | | | |
QY 2101 AAACACACCCGCGCAGGCAAGCCGACTTCAGGGAATGAGGAGCTCCGAGCCGTTTGGCAGAGATC 2160
| | | | |
Db 2437 AAACACACCCGCGCAGGCAAGCCGACTTCAGGGAATGAGGAGCTCCGAGCCGTTTGGCAGAGATC 2496
| | | | |
QY 2161 ACGGAGATGTGTGACACAGCTCCAGGTAACGATGATCAAGATGGAATTTCCAGAAAGCTG 2220
| | | | |
Db 2497 ACGGAGATGTGTGACACAGCTCCAGGTAACGATGATCAAGATGGAATTTCCAGAAAGCTG 2556
| | | | |
QY 2221 CACGAATCAAGAAAGATTGATGATGGAATTCATCTTGTGTTCCGGAAGAGGAGTTTC 2280
| | | | |
Db 2557 CACGAATCAAGAAAGATTGATGATGGAATTCATCTTGTGTTCCGGAAGAGGAGTTTC 2616
| | | | |
QY 2281 ATCCGTCTGGGAGGCTCAGCAAGCTCTCGGGGAAGGGGCTTCACAGCGCATGTTCTTC 2340
| | | | |
Db 2617 ATCCGTCTGGGAGGCTCAGCAAGCTCTCGGGGAAGGGGCTTCACAGCGCATGTTCTTC 2676
| | | | |
QY 2341 CTGTTCAACGACGTCTGCTATACACGAGCCGGGGGCTGACGGCTTCCATTAAGTTTAA 2400
| | | | |
Db 2677 CTGTTCAACGACGTCTGCTATACACGAGCCGGGGGCTGACGGCTTCCATTAAGTTTAA 2736
| | | | |
QY 2401 GTCCAGGGGAGCTCCGCTCTATGGCATGACGATTTGAGAGAGCGAAGACGAGTGGGG 2460
| | | | |
Db 2737 GTCCAGGGGAGCTCCGCTCTATGGCATGACGATTTGAGAGAGCGAAGAGTGGGG 2796
| | | | |
QY 2461 GTGCCCACTGCTGAACCTCCCGGGCCAGCGGAGTCCATCATCTGTGGCGCCAGATTCT 2520
| | | | |
Db 2797 GTGCCCACTGCTGAACCTCCCGGGCCAGCGGAGTCCATCATCTGTGGCGCCAGATTCT 2856
| | | | |
QY 2521 CGGTCCGAGATGAGAAATGTGGTTGAGAGACATCCAGATGGCCATTGACCTGGCGGAAGAG 2580
| | | | |
Db 2857 CGGTCCGAGATGAGAAATGTGGTTGAGAGACATCCAGATGGCCATTGACCTGGCGGAAGAG 2916
| | | | |
QY 2581 AGCAGACGCCCGCCCTGAGTCTCTGAGCCAGACGCCCTTGAACAACAAGTCCCTGAT 2640
| | | | |
Db 2917 AGCAGACGCCCGCCCTGAGTCTCTGAGCCAGACGCCCTTGAACAACAAGTCCCTGAT 2976
| | | | |
QY 2641 GAAGCCACCGCGGCTGACCAAGAGTCAAGAGATGACTTGAAGCGCTCGCGCATCTGCTG 2700
| | | | |
Db 2977 GAAGCCACCGCGGCTGACCAAGAGTCAAGAGATGACTTGAAGCGCTCGCGCATCTGCTG 3036
| | | | |
QY 2701 GAGGCCCAGGCCCGGCAACCGGGCAACATATGTGTGACGTGTGTGCAACCGCAACACC 2760
| | | | |

DB 3037 GAGCGCCAGGCGCCGCGCACCGGCAACACATGATGTCGTGTCGACCGCAACACC 3096
QY 2761 AGGCTCTCCATGGTGAATTCACATCGCATGGAGAAATCAATGTTGTTGAAAACCTGCTG 2820
DB 3097 AGGCTCTCCATGGTGAATTCACATCGCATGGAGAAATCAATGTTGTTGAAAACCTGCTG 3156
QY 2821 AGGAAATTCAAAACAGCAACGGGTGGCAGAACTGTGGTGTGTTCACAAACTTCTG 2880
DB 3157 AGGAAATTCAAAACAGCAACGGGTGGCAGAACTGTGGTGTGTTCACAAACTTCTG 3216
QY 2881 CTGTTCTTTCACAAATCACACAGAACATCATCCCTTCCAGACCTGCTGCTGCGC 2940
DB 3217 CTGTTCTTTCACAAATCACACAGAACATCATCCCTTCCAGACCTGCTGCTGCGC 3276
QY 2941 TACTGCTCACCATCCCTCTGTGATCCGAGAAATCCAGAAACATTCAGTTCAAGCTG 3000
DB 3277 TACTGCTCACCATCCCTCTGTGATCCGAGAAATCCAGAAACATTCAGTTCAAGCTG 3336
QY 3001 CACTTCAAGTCCGACGCTCTACTACTTCAAGGCGGAAAACGAGTACAGTTCAAGGCTG 3060
DB 3337 CACTTCAAGTCCGACGCTCTACTACTTCAAGGCGGAAAACGAGTACAGTTCAAGGCTG 3396
QY 3061 ATGGAAGTATCCGACGATGCGACAGCTCTGCTCGGACCCGACGTTGAGCCACAA 3120
DB 3397 ATGGAAGTATCCGACGATGCGACAGCTCTGCTCGGACCCGACGTTGAGCTACAA 3456
QY 3121 GAGTCTCTTGTGTAT 3135
DB 3457 GAGTCTCTTGTGTAT 3471
RESULT 5
BD231200 4687 bp DNA linear PART 17-JUL-2003
LOCUS Human cytochrome associated proteins.
DEFINITION BD231200
ACCESSION BD231200.1 GI:33040970
VERSION JP 2002526076-A/2.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
1 (bases 1 to 4687)
Lal,P., Tang,T.Y., Yue,H., Hillman,J.L., Bandman,O., Corley,N.C.,
Guegler,K.J., Patterson,C., Azimzai,Y. and Baughn,M.R.
Human cytochrome associated proteins
Patent: JP 2002526076-A 2 20-AUG-2002;
INCYTE PHARMACEUTICALS INC
OS Homo sapiens (human)
PN JP 2002526076-A/2
PD 20-AUG-2002
PR 17-SEP-1999 JP 2000574254
PR 18-SEP-1998 US 60/172226,27-APR-1999 US 60/131321 PI
PRESET LAL,TOM Y TANG,HENRY YUE,JENNIFER L HILLMAN,OLGA PI
BANDMAN
PI NEIL C CORLEY,KARL J GUEGLER,CHANDRA PATTERSON,YALDA AZIMZAI,
PI MARIAH R BAUGHN
PC C12N15/09,A61K38/00,A61K45/00,A61P1/16,A61P3/00,A61P3/06 PC
A61P3/10,A61P5/00,
PC A61P5/38,A61P7/00,A61P7/06,A61P9/00,A61P9/10,A61P11/06 PC
A61P13/00,A61P13/12,
PC A61P13/06,A61P19/02,A61P19/10,A61P21/04,A61P25/00,A61P25/00,
PC A61P25/04,
PC A61P25/14,A61P25/16,A61P25/28,A61P29/00,A61P31/04,A61P31/12,
PC A61P31/18,
PC A61P33/00,A61P35/00,A61P35/02,A61P37/00,A61P43/00,C07K14/47,
PC C07K16/18,
PC C12N1/15,C12N1/19,C12N1/21,C12N5/10,C12P21/02,C12P22/08 PC
C12O1/68,G01N33/15,
PC G01N33/50,G01N33/53,G01N33/53,G01N33/566,C12N15/00,C12N5/00,
PC A61K37/02
CC Incyte ID No: 1320252

PH Key Location/Qualifiers
FT source 1..4687
/organism='Homo sapiens (human)'.
FEATURES
source Location/Qualifiers
1..4687
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 99.9%; Score 3131.8; DB 2; Length 4687;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3133; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATGGAGAAATAGAGAGAGGCGGACCCAGAGATCAAGCTGGGGGCCCCGGAAATTTGG 60
DB 90 ATGGAGAAATAGAGAGAGGCGGACCCAGAGATCAAGCTGGGGGCCCCGGAAATTTGG 149
QY 61 GGGATCAGTACCTTGGAGAGTGGACAGAAAGCCGCCCAACACTTCAGGAAATCGTG 120
DB 150 GGGATCAGTACCTTGGAGAGTGGACAGAAAGCCGCCCAACACTTCAGGAAATCGTG 209
QY 121 TCCATCAAAATCCAGATGCTGATGACACCGAGAGGCAATTTGAATTCACAAAGAGCT 180
DB 210 TCCATCAAAATCCAGATGCTGATGACACCGAGAGGCAATTTGAATTCACAAAGAGCT 269
QY 181 CTTGGGAGAGTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
DB 270 CTTGGGAGAGTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 329
QY 241 TTTGGCTGAGTTCTGATCAAAAGATCAAGGATGATGATGATGATGATGATGATGATGAT 300
DB 330 TTTGGCTGAGTTCTGATCAAAAGATCAAGGATGATGATGATGATGATGATGATGATGAT 389
QY 301 ATTTGAAAACAGATTAGAAAGCCAAAGCACTGTTGTTTAAAGTTTGTGTAATTTCTTT 360
DB 390 ATTTGAAAACAGATTAGAAAGCCAAAGCACTGTTGTTTAAAGTTTGTGTAATTTCTTT 449
QY 361 CGGCTGACACACACACTTCCAAAGAAATTCACAAAGTACTGTTGCGCTGCAAGGTG 420
DB 450 CGGCTGACACACACACTTCCAAAGAAATTCACAAAGTACTGTTGCGCTGCAAGGTG 509
QY 421 AAGCAGACTTGGCTCAAGGAGGTTGACGTGTAATGACACGAGGACAGCTCTTGATTT 480
DB 510 AAGCAGACTTGGCTCAAGGAGGTTGACGTGTAATGACACGAGGACAGCTCTTGATTT 569
QY 481 TCACACATTGTCAATCTGAGATTGGGATTTTGTATGAAGCTTGGACAGAGCACTTA 540
DB 570 TCACACATTGTCAATCTGAGATTGGGATTTTGTATGAAGCTTGGACAGAGCACTTA 629
QY 541 GCAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 600
DB 630 GCAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 689
QY 601 CATTAACCAATTGACAAACACAGAGAAATGAGATTTCAGCTCCTTAGAGTTGCCGT 660
DB 690 CATTAACCAATTGACAAACACAGAGAAATGAGATTTCAGCTCCTTAGAGTTGCCGT 749
QY 661 CGGCTAGAGATGATGAAATCCGTTGCAACCGGCGCAAGAGAGAGAGAGAGAGAGATC 720
DB 750 CGGCTAGAGATGATGAAATCCGTTGCAACCGGCGCAAGAGAGAGAGAGAGAGATC 809
QY 721 AATTCGCGCTGGCCCAACCGGAAATTCAGTGTTCAGGGTTTCACTAAGATTCATGTC 780
DB 810 AATTCGCGCTGGCCCAACCGGAAATTCAGTGTTCAGGGTTTCACTAAGATTCATGTC 869
QY 781 TTCACCTGGGCGCAAGGTGGAGAGCTGAGCTTCAAGAGAGAGAGAGAGAGAGAGT 840
DB 870 TTCACCTGGGCGCAAGGTGGAGAGCTGAGCTTCAAGAGAGAGAGAGAGAGAGT 929
QY 841 CGGCGAGATGCAATAGTGCATACAGAGATACCTTGAATTCCTGATGCGAGTGGAGT 900
DB 930 CGGCGAGATGCAATAGTGCATACAGAGATACCTTGAATTCCTGATGCGAGTGGAGT 989

QY 901 TTCTGCAAGTCTTCTGAAAAATCTGTGTGAACATCATGCTTCTTTAGACTTTTGA 960
| | | | |
Db 990 TTCTGCAATCTCTTCTGAAAAATCTGTGTGAACATCATGCTTCTTTAGACTTTTGA 1049
| | | | |
QY 961 GAGCCCAACCAAAAGCCCAAGCCCGTCTCTTTAGCCGGGGGTCAATCATTTGGTTCACT 1020
| | | | |
Db 1050 GAGCCCAACCAAAAGCCCAAGCCCGTCTCTTTAGCCGGGGGTCAATCATTTGGTTCACT 1109
| | | | |
QY 1021 GGTCCGACCTCAGAAAGAGGTTCTCGACTATGTATAAGAGAGACATAGAAAGGTGCAG 1080
| | | | |
Db 1110 GGTCCGACCTCAGAAAGAGGTTCTCGACTATGTATAAGAGAGACATAGAAAGGTGCAG 1169
| | | | |
QY 1081 TTTGAAAGGAGACACAGCAAGATTCTATTCGGAGCTTGCTTACAGCTACAGAA 1140
| | | | |
Db 1170 TTTGAAAGGAGACACAGCAAGATTCTATTCGGAGCTTGCTTACAGCTACAGAA 1229
| | | | |
QY 1141 CTGAATTCGGAAGTGTCTGAGACAGTCTCAGCAGAGACCAAGCTTTACATTTGAGAAAGT 1200
| | | | |
Db 1230 CTGAATTCGGAAGTGTCTGAGACAGTCTCAGCAGAGACCAAGCTTTACATTTGAGAAAGT 1289
| | | | |
QY 1201 GCCGAATCTTCAGGGGGCCAGAGCTGCGCGGAGAAAGAAACCGAAGTTTCCGCGGG 1260
| | | | |
Db 1290 GCCGAATCTTCAGGGGGCCAGAGCTGCGCGGAGAAAGAAACCGAAGTTTCCGCGGG 1349
| | | | |
QY 1261 GAGCCGGGGTCCACCCGAGACCTTGCAGGAGAGAAAGCCCGGGGTAAACAAGCAGCG 1320
| | | | |
Db 1350 GAGCCGGGGTCCACCCGAGACCTTGCAGGAGAGAAAGCCCGGGGTAAACAAGCAGCG 1409
| | | | |
QY 1321 GACGAGACCGGCTCGGCGCCACGAGAGAAAGAGAGAGTCTTTAAGATAGAGCCAG 1380
| | | | |
Db 1410 GACGAGACCGGCTCGGCGCCACGAGAGAAAGAGAGAGTCTTTAAGATAGAGCCAG 1469
| | | | |
QY 1381 CAGAGTAACTCAGGCCCCCGAGCCCAAGCAAGGCTCCCTGACCTGGAGTCTCAACT 1440
| | | | |
Db 1470 CAGAGTAACTCAGGCCCCCGAGCCCAAGCAAGGCTCCCTGACCTGGAGTCTCAACT 1529
| | | | |
QY 1441 TCCGAGCTGTCTGTGAACCTCGAGGGGGAGTGGCCCTGCAACCTGACTTGTCTCC 1500
| | | | |
Db 1530 TCCGAGCTGTCTGTGAACCTCGAGGGGGAGTGGCCCTGCAACCTGACTTGTCTCC 1589
| | | | |
QY 1501 AACCTGAGCCCGGACACCAAGCAGGCTCTCCCTTGATCAGCCCGTGTGATAGACAG 1560
| | | | |
Db 1590 AACCTGAGCCCGGACACCAAGCAGGCTCTCCCTTGATCAGCCCGTGTGATAGACAG 1649
| | | | |
QY 1561 GCCTGCCCCGAGCAGATGAGATGAGGGCCGGAGAGAGATTCCCACTGTATA 1620
| | | | |
Db 1650 GCCTGCCCCGAGCAGATGAGATGAGGGCCGGAGAGAGATTCCCACTGTATA 1709
| | | | |
QY 1621 GCGTACTTATAGCTAAGGAAGTGTCTACACCGAGCAGAACATATCTGAAGATCTCGAA 1680
| | | | |
Db 1710 GCGTACTTATAGCTAAGGAAGTGTCTACACCGAGCAGAACATATCTGAAGATCTCGAA 1769
| | | | |
QY 1681 GTTATACCTTCTGTGTTTCAAGCAGATGAGCAAAAGAGAGCCATGCGCGAGACATG 1740
| | | | |
Db 1770 GTTATACCTTCTGTGTTTCAAGCAGATGAGCAAAAGAGAGCCATGCGCGAGACATG 1829
| | | | |
QY 1741 AAAAGTCTATATTCGGAATTTTGAACCTTTGCAAAATTTCTATCTAATTTTCTCAG 1800
| | | | |
Db 1830 AAAAGTCTATATTCGGAATTTTGAACCTTTGCAAAATTTCTAATTTTCTCAG 1889
| | | | |
QY 1801 GAAATTTGAGCAAGACTTGCCTGTGGAGAGGCGCTCAATATGCCAAATCTAGAGATTAC 1860
| | | | |
Db 1890 GAAATTTGAGCAAGACTTGCCTGTGGAGAGGCGCTCAATATGCCAAATCTAGAGATTAC 1949
| | | | |
QY 1861 CAAAGATGCGGAGATGTCACTGCTGAAGAACTTCAAGGCAATGAGACCTGGCGGTCAAC 1920
| | | | |
Db 1950 CAAAGATGCGGAGATGTCACTGCTGAAGAACTTCAAGGCAATGAGACCTGGCGGTCAAC 2009
| | | | |
QY 1921 CTGTGGAACACAGGAGGCTTTGAGGCGCTGGAAGATGGAATCAAGAGCTCCGCGCG 1980
| | | | |
Db 2010 CTGTGGAACACAGGAGGCTTTGAGGCGCTGGAAGATGGAATCAAGAGCTCCGCGCG 2069
| | | | |

QY 1981 CTGGAAGACTTCTGAGAGACTTTGAGCTGAGAGAGTGTGTATCAACCGCTCAACAC 2040
| | | | |
Db 2070 CTGGAAGACTTCTGAGAGACTTTGAGCTGAGAGAGTGTGTATCAACCGCTCAACAC 2129
| | | | |
QY 2041 TTCTCTCTGCGGCACTGACACCGGCTCATGTCACTACAGAGGTCTTGAAGGGCTGTGC 2100
| | | | |
Db 2130 TTCTCTCTGCGGCACTGACACCGGCTCATGTCACTACAGAGGTCTTGAAGGGCTGTGC 2189
| | | | |
QY 2101 AAACACACCCGCGAGGACGCGGACTTTCAGGGACTGCGGAGCGCTTTTGGCAGAGATC 2160
| | | | |
Db 2190 AAACACACCCGCGAGGACGCGGACTTTCAGGGACTGCGGAGCGCTTTTGGCAGAGATC 2249
| | | | |
QY 2161 ACGAGATGTGAGCAAGCTCCACGGTACATGATCAAGATGAGAAATTTCCAGAGCTG 2220
| | | | |
Db 2250 ACGAGATGTGAGCAAGCTCCACGGTACATGATCAAGATGAGAAATTTCCAGAGCTG 2309
| | | | |
QY 2221 CACGAATCTCAAGAAAGATTTGATTTGGCATTTGTGATTTTCCGGAGAGGAGTTT 2280
| | | | |
Db 2310 CACGAATCTCAAGAAAGATTTGATTTGGCATTTGTGATTTTCCGGAGAGGAGTTT 2369
| | | | |
QY 2281 ATCGCTGAGGAGCTCAGCAAGCTCTCGGGGAAAGGGGCTTCAGCAGCGCATTTCTTC 2340
| | | | |
Db 2370 ATCGCTGAGGAGCTCAGCAAGCTCTCGGGGAAAGGGGCTTCAGCAGCGCATTTCTTC 2429
| | | | |
QY 2341 CTGTTCAACGAGGCTCTGCTATACAGAGCCGGGGGCTGACGGCTTCAATCATGTTTAA 2400
| | | | |
Db 2430 CTGTTCAACGAGGCTCTGCTATACAGAGCCGGGGGCTGACGGCTTCAATCATGTTTAA 2489
| | | | |
QY 2401 GTTCAAGGGCAGCTCCCGCTCTATGAGCATGACATTTAGAGAGAGGAAAGACAGAGTGGGG 2460
| | | | |
Db 2490 GTTCAAGGGCAGCTCCCGCTCTATGAGCATGACATTTAGAGAGAGGAAAGACAGAGTGGGG 2549
| | | | |
QY 2461 GTGCCCACTGCTGACCTCCGAGGGCAGCGGAGTCCATCATCTGTGGCCGCACTTCT 2520
| | | | |
Db 2550 GTGCCCACTGCTGACCTCCGAGGGCAGCGGAGTCCATCATCTGTGGCCGCACTTCT 2609
| | | | |
QY 2521 CGGTCCAGATGAGAGAGTGGGTTGAGAACATCCAGATGGCCATTGACCTGGCGGAGAG 2580
| | | | |
Db 2610 CGGTCCAGATGAGAGAGTGGGTTGAGAACATCCAGATGGCCATTGACCTGGCGGAGAG 2669
| | | | |
QY 2581 AGCAGACACCCCGCCCTGAGTCTCGGCGCAGAGCCCCCTGCAACAAGTCCCTGAT 2640
| | | | |
Db 2670 AGCAGACACCCCGCCCTGAGTCTCGGCGCAGAGCCCCCTGCAACAAGTCCCTGAT 2729
| | | | |
QY 2641 GAAAGCACCAGGCTGACACAGAGTCAAGAGATGACTGAGGCGCTCCGCGCATCTGCTG 2700
| | | | |
Db 2730 GAAAGCACCAGGCTGACACAGAGTCAAGAGATGACTGAGGCGCTCCGCGCATCTGCTG 2789
| | | | |
QY 2701 GAGCGCCAGGCCCCGACCCGCGGCAACAATGCTGACAGTGTGCGACCCGCAACACC 2760
| | | | |
Db 2790 GAGCGCCAGGCCCCGACCCGCGGCAACAATGCTGACAGTGTGCGACCCGCAACACC 2849
| | | | |
QY 2761 AGCGTCTCATAGTGGACTTTCAGCATGCGAGTGGAGATTCAGTTGTCTGAAACCTGCTG 2820
| | | | |
Db 2850 AGCGTCTCATAGTGGACTTTCAGCATGCGAGTGGAGATTCAGTTGTCTGAAACCTGCTG 2909
| | | | |
QY 2821 AGGAATTTCAAAAAACAGCAAGGCTGAGAGAGTGTGGTGTGTTCACAACTTTCTGC 2880
| | | | |
Db 2910 AGGAATTTCAAAAAACAGCAAGGCTGAGAGAGTGTGGTGTGTTCACAACTTTCTGC 2969
| | | | |
QY 2881 CTGTTCTTCTACAAATCAACACAGCAATATCCCTTTGCCAGCTGCTGCTCGGC 2940
| | | | |
Db 2970 CTGTTCTTCTACAAATCAACACAGCAATATCCCTTTGCCAGCTGCTGCTCGGC 3029
| | | | |
QY 2941 TACTGCTCAACATCCCTGAGTCCGAGAACATCCAGAAAGCTAGTGTTCAGAGCTG 3000
| | | | |
Db 3030 TACTGCTCAACATCCCTGAGTCCGAGAACATCCAGAAAGCTAGTGTTCAGAGCTG 3089
| | | | |
QY 3001 CACTTCAAGTCCACAGTCTACTACTTTCAGGGCGGAAACGAGTACAGCTTTCAGAAAGTGG 3060
| | | | |
Db 3090 CACTTCAAGTCCACAGTCTACTACTTTCAGGGCGGAAACGAGTACAGCTTTCAGAAAGTGG 3149
| | | | |
QY 3061 ATGGAAGTATCCGAGTGGCCACAGCTCTGCGGACCCCAAGTGTTCAGCCACAA 3120
| | | | |

Db 3150 ATGGAAGTATCCGACATGCGACAGCTGCGCGGACCCGACCGTGTGATGATCAAA 3209
Qy 3121 GAGTCTCTTGTGTAT 3135
Db 3210 GAGTCTCTTGTGTAT 3224

RESULT 6
BC071592
LOCUS BC071592 4932 bp mRNA linear PRI 28-JUL-2005
DEFINITION Homo sapiens FERM, RhogEF (ARHGEF) and plecterin domain protein 1 (chondrocyte-derived), mRNA (cDNA clone MGC:87400 IMAGE:30342239), complete cds.
BC071592
VERSION BC071592.1 GI:47940451
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 4932)
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Haide, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Uedln, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Viatello, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hu, X., Gibbs, R.A., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Hellon, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butlerfield, Y.S., Krzywinski, M.I., Skalski, U., Smalins, D.E., Schermer, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Mammalian Gene Collection Program Team
Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 4932)
NIH MGC Project
Direct Submission
Submitted (01-JUN-2004) National Institutes of Health, Mammalian Gene Collection (MGC), Bethesda, MD 20892-2590, USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: Dr. Stefan Hanson
cDNA Library Preparation: Michael Brownstein / Ted Usdin
Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-ehgc.stanford.edu>
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNL at: <http://image.llnl.gov>
Series: IRAP Plate: 168 Row: 1 Column: 8
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 48928036.
Location/Qualifiers
1. .4932

FEATURES
source

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:87400 IMAGE:30342239"
/issue_type="Placenta, normal"
/clone_lib="NIH MGC_147"
/lab_host="DH10B"
/note="Vector: pBluescriptCR"
1. .4932
/gene="FARP1"
/note="synonyms: CDEP, MGC87400, PLEKHG2"
/db_xref="GeneID:10160"
/db_xref="MIM:602654"
240_ .3470
/gene="FARP1"
/codon_start=1
/product="FARP1 protein"
/protein_id="AAH71592.1"
/db_xref="GI:47940452"
/db_xref="GeneID:10160"
/db_xref="MIM:602654"
CDs
gene
CDs
ORIGIN
Query Match 96.6%; Score 3028.8; DB 5; Length 4932;
Best Local Similarity 97.1%; Pred. No. 0;
Matches 3133; Conservative 0; Mismatches 2; Indels 93; Gaps 1;
Qy 1 ATGGAGAATATAGACAGAGCCGACCCGACGATCAAGCTGGGGGCCCCGGAATTCG 60
Db 240 ATGGAGAATATAGACAGAGCCGACCCGACGATCAAGCTGGGGGCCCCGGAATTCG 239
Qy 61 GGGATCAGTACCTTGGAAAGTGGACGAAGCCGCCCAACACCTTCAGGAAATCTCGTG 120
Db 300 GGGATCAGTACCTTGGAAAGTGGACGAAGCCGCCCAACACCTTCAGGAAATCTCGTG 359
Qy 121 TCCATCAAAATCCAGATGCTGATGACACCCAGAGGACATTTGAAGTTCCAAAGAGCT 180
Db 360 TCCATCAAAATCCAGATGCTGATGACACCCAGAGGACATTTGAAGTTCCAAAGAGCT 419
Qy 181 CTTGGAGAAGTGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
Db 420 CTTGGAGAAGTGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 479
Qy 241 TTTGGCTGAGTTTCTGATCAAAAAGATGATGATGATGATGATGATGATGATGATGAT 300
Db 480 TTTGGCTGAGTTTCTGATCAAAAAGATGATGATGATGATGATGATGATGATGATGAT 539
Qy 301 ATTGTAACAGATTGAAGGCCAAGCAAGCTGTTGTTGAAGTTTGGTGAATTCCTT 360
Db 540 ATTGTAACAGATTGAAGGCCAAGCAAGCTGTTGTTGAAGTTTGGTGAATTCCTT 599
Qy 361 CCGCTGACCAACACACTCCAAAGAACTCACAAGTACCTGTTGCGCGTGCAGGTG 420
Db 600 CCGCTGACCAACACACTCCAAAGAACTCACAAGTACCTGTTGCGCGTGCAGGTG 659

OY 421 AAGCAGAGCTTGGCTCAAGCAGGTGAGTGTATGACACGACGACGCTCTTGATT 480
| | | | |
DB 660 AAGCAGAGCTTGGCTCAAGCAGGTGAGTGTATGACACGACGACGCTCTTGATT 719
| | | | |
OY 481 TCACACATTTGTGCAATCTGAGATTGGGAAATTTTGTATGAAGCTTTGACAGAGACCTTA 540
| | | | |
DB 720 TCACACATTTGTGCAATCTGAGATTGGGAAATTTTGTATGAAGCTTTGACAGAGACCTTA 779
| | | | |
OY 541 GCAAAAATTAATTAATCTTACTCTGAGCAAGCGCATGAGAGCAAAATCCGTGAAATTTTCA 600
| | | | |
DB 780 GCAAAAATTAATTAATCTTACTCTGAGCAAGCGCATGAGAGCAAAATCCGTGAAATTTTCA 839
| | | | |
OY 601 CATTAACCATTTGAGCAAAACACGAGCAAGATCAGATTTTCCAGCTCTTAAGATTTGCCGT 660
| | | | |
DB 840 CATTAACCATTTGAGCAAAACACGAGCAAGATCAGATTTTCCAGCTCTTAAGATTTGCCGT 899
| | | | |
OY 661 CCGCTAGAGATGTATGAAATCCGGTTGCACCCGGCCAAAGACAGGGAAAGCCAGAAAGATC 720
| | | | |
DB 900 CCGCTAGAGATGTATGAAATCCGGTTGCACCCGGCCAAAGACAGGGAAAGCCAGAAAGATC 959
| | | | |
OY 721 AATCTGGCCGTTGGCCCAACGGGAATTTCTAGTGTTCAGGGTTTCACTAAGATCAATGCG 780
| | | | |
DB 960 AATCTGGCCGTTGGCCCAACGGGAATTTCTAGTGTTCAGGGTTTCACTAAGATCAATGCG 1019
| | | | |
OY 781 TTCAACTGGGCCCAGAGTGCGGAAGCTGAGCTTCAAGAGAAAGCGCTTCTCAATCAAGCTC 840
| | | | |
DB 1020 TTCAACTGGGCCCAGAGTGCGGAAGCTGAGCTTCAAGAGAAAGCGCTTCTCAATCAAGCTC 1079
| | | | |
OY 841 CCGCCAGATGCCAATATGTGCGTACAGAGATACCTTGAAATTCCTGATGCGCAGTCCGGAT 900
| | | | |
DB 1080 CCGCCAGATGCCAATATGTGCGTACAGAGATACCTTGAAATTCCTGATGCGCAGTCCGGAT 1139
| | | | |
OY 901 TTCTGAAATGCTCTTGGAAAAATCTGTGTGAACATCATGCTCTTGAACCTTTTGA 960
| | | | |
DB 1140 TTCTGAAATGCTCTTGGAAAAATCTGTGTGAACATCATGCTCTTGAACCTTTTGA 1199
| | | | |
OY 961 GAGCCCAAAACAAAGGCCCAAGCCGCTCTTTAGCCGGGGGTCACTATTTGCGTTCAGT 1020
| | | | |
DB 1200 GAGCCCAAAACAAAGGCCCAAGCCGCTCTTTAGCCGGGGGTCACTATTTGCGTTCAGT 1259
| | | | |
OY 1021 GGTGCGACTCAGAGCAAGGTTCTGACTATGTTAAAGAGAGACATTAAGAGGTGCGAG 1080
| | | | |
DB 1260 GGTGCGACTCAGAGCAAGGTTCTGACTATGTTAAAGAGAGACATTAAGAGGTGCGAG 1319
| | | | |
OY 1081 TTGGAAGAGAGACAGCAAGATTCAATTCATCCGAGCTTGCTTCAAGCTTAAGAA 1140
| | | | |
DB 1320 TTGGAAGAGAGACAGCAAGATTCAATTCATCCGAGCTTGCTTCAAGCTTAAGAA 1379
| | | | |
OY 1141 CTGAATTCGGAAGTGTCTGAGCAGTCTCAGCAGAGACCAAGCCTTACATTTGAGAGAGT 1200
| | | | |
DB 1380 CTGAATTCGGAAGTGTCTGAGCAGTCTCAGCAGAGACCAAGCCTTACATTTGAGAGAGT 1439
| | | | |
OY 1201 GCCGAAATCTCAAGGGGGCCAGAGCTCCGGCGAGGAAGAACCGAAGGTTTCCGCGAG 1260
| | | | |
DB 1440 GCCGAAATCTCAAGGGGGCCAGAGCTCCGGCGAGGAAGAACCGAAGGTTTCCGCGAG 1499
| | | | |
OY 1261 GAGCCGGGGGTGCGACCGAGCCCTGCGCGAGAGGAAGCCCGCGGGTTAAACAAGCAGGGG 1320
| | | | |
DB 1500 GAGCCGGGGGTGCGACCGAGCCCTGCGCGAGAGGAAGCCCGCGGGTTAAACAAGCAGGGG 1559
| | | | |
OY 1321 GACGAGCCGCTCGCGGCCCAAGAGGAAGAGAGAGAGTCTGTTAAGATAGAACCCAG 1380
| | | | |
DB 1560 GACGAGCCGCTCGCGGCCCAAGAGGAAGAGAGAGAGTCTGTTAAGATAGAACCCAG 1619
| | | | |
OY 1381 CAGAGTAAATCTAGCCCCCGCAGCCAGACAGAGCTCCCTGACTGGCAGTCTCACTT 1440
| | | | |
DB 1620 CAGAGTAAATCTAGCCCCCGCAGCCAGACAGAGCTCCCTGACTGGCAGTCTCACTT 1679
| | | | |
OY 1441 TCCGAGCTGTCTGTGAATCTGCAAGGGGGAGTGGCCCTGSCAAAGTGAACCTTGTCTCC 1500
| | | | |
DB 1680 TCCGAGCTGTCTGTGAATCTGCAAGGGGGAGTGGCCCTGSCAAAGTGAACCTTGTCTCC 1739
| | | | |

OY 1501 AACCTGAGCCCGACACCAAGCAGGCTCTCCCTTGATCAGCCGCTGTGTATGACAG 1560
| | | | |
DB 1740 AACCTGAGCCCGACACCAAGCAGGCTCTCCCTTGATCAGCCGCTGTGTATGACAG 1799
| | | | |
OY 1561 GCGTGGCCCCCGGACCGAGATGAGAGTGGGCCGAGAGAGATTTCCCACTGATTA 1620
| | | | |
DB 1800 GCGTGGCCCCCGGACCGAGATGAGAGTGGGCCGAGAGAGATTTCCCACTGATTA 1859
| | | | |
OY 1621 GCGTACTCATGTAGCTAAGAAATGTCTACACCGAGCGCAATATCTGAAAGATCTGAA 1680
| | | | |
DB 1860 GCGTACTCATGTAGCTAAGAAATGTCTACACCGAGCGCAATATCTGAAAGATCTGAA 1919
| | | | |
OY 1681 GTTATCACTTGTGTGTTTCAAGACAGTGAACAAAGAGAGCGCATGCGGAGCACTG 1740
| | | | |
DB 1920 GTTATCACTTGTGTGTTTCAAGACAGTGAACAAAGAGAGCGCATGCGGAGCACTG 1979
| | | | |
OY 1741 AAAAGTCTCATTTTCCCGAATTTTGAACCTTTGCAAAATTTTCAATTAATTTTCTCAAG 1800
| | | | |
DB 1980 AAAAGTCTCATTTTCCCGAATTTTGAACCTTTGCAAAATTTTCAATTAATTTTCTCAAG 2039
| | | | |
OY 1801 GAAATTTGAGCAACGACTTGCCCTGTGGGAAGGCGCTCAATGCCCCAAATCAGAGATTAC 1860
| | | | |
DB 2040 GAAATTTGAGCAACGACTTGCCCTGTGGGAAGGCGCTCAATGCCCCAAATCAGAGATTAC 2099
| | | | |
OY 1861 CAAGAATCGCGGATGTCACTGTGAAGAACTTCAAGGCGATGAACACCTGGCGGCTCAC 1920
| | | | |
DB 2100 CAAGAATCGCGGATGTCACTGTGAAGAACTTCAAGGCGATGAACACCTGGCGGCTCAC 2159
| | | | |
OY 1921 CTGTGGAAGCACAGGAGGCTTTGAGAGCCCTTGGAGAAATGGAATCAAGAGCTCCGGCGG 1980
| | | | |
DB 2160 CTGTGGAAGTACAGGAGGCTTTGAGAGCCCTTGGAGAAATGGAATCAAGAGCTCCGGCGG 2219
| | | | |
OY 1981 CTGGAAGACTTCTGAGAGACTTTGAGCTGCAAGAGTGTGTAACTACCTCCGCTCAACAC 2040
| | | | |
DB 2220 CTGGAAGACTTCTGAGAGACTTTGAGCTGCAAGAGTGTGTAACTACCTCCGCTCAACAC 2279
| | | | |
OY 2041 TTCTCTCTGCGGCACTGACCGGCTCATGCACTAACAGAGTCTTGAGCGGCTGTGTC 2100
| | | | |
DB 2280 TTCTCTCTGCGGCACTGACCGGCTCATGCACTAACAGAGTCTTGAGCGGCTGTGTC 2339
| | | | |
OY 2101 AAAACACACCGCCGAGCCAGCCGACCTTCAAGGACTGCGGACCGCTTTGGCAGAGATC 2160
| | | | |
DB 2340 AAAACACACCGCCGAGCCAGCCGACCTTCAAGGACTGCGGACCGCTTTGGCAGAGATC 2399
| | | | |
OY 2161 ACGAGATGTGTGACACAGCTCCACGCTACAGATGATCAAGATGAGAAATTTCCAGAAAGCTG 2220
| | | | |
DB 2400 ACGAGATGTGTGACACAGCTCCACGCTACAGATGATCAAGATGAGAAATTTCCAGAAAGCTG 2459
| | | | |
OY 2221 CACGAATCTCAAGAAAGATTGATGTCATTTGTGTTCCGGGAAG----- 2273
| | | | |
DB 2460 CACGAATCTCAAGAAAGATTGATGTCATTTGTGTTCCGGGAAGCGCTGGC 2519
| | | | |
OY 2274 ----- 2273
| | | | |
DB 2520 TCCTTTTCCCTTATGAGGACTTCACACTTAAGACCAAGCGCATCCCATGTGCTCA 2579
| | | | |
OY 2274 -----GGAGTTTCATCCGCTTGGGAGCCTCAGCAAGCTC 2307
| | | | |
DB 2580 GAGAGAGGCCATTGTTATTAAGAGAGTTCACTCGCTGTGGAGCCTCAGCAAGCTC 2639
| | | | |
OY 2308 TCGGGGAAGGGGCTCCAGAGCGCATGTTCTTCTGTTCACAGAGCTCTGCTATACAG 2367
| | | | |
DB 2640 TCGGGGAAGGGGCTCCAGAGCGCATGTTCTTCTGTTCACAGAGCTCTGCTATACAG 2699
| | | | |
OY 2368 AGCCGGGGGCTGACGGGCTTCCAACTAGTTTAAAGTCAACGGGCACTCCCGCTTAATGGC 2427
| | | | |
DB 2700 AGCCGGGGGCTGACGGGCTTCCAACTAGTTTAAAGTCAACGGGCACTCCCGCTTAATGGC 2759
| | | | |
OY 2428 ATGAGATTTGAGAGAGGAGACAGAGTGGGGGTGGCCCACTGCTGACCTCTCGGGGC 2487
| | | | |
DB 2760 ATGAGATTTGAGAGAGGAGACAGAGTGGGGGTGGCCCACTGCTGACCTCTCGGGGC 2819
| | | | |
OY 2488 CAGCGGCAAGTCCATCATGTGTGCGCAGTTCTCGGTCGAGATGAGAAAGTGGGTTGAG 2547
| | | | |

```

Db      2820 CAGGGGAGTTCATCATCTGTGGCCGAGTTCTCGTCCGAGATGAGGAAGTGGTTGAG 2879
Qy      2548 GACATCCAGATGGCCATTGACCTTGGCGAGAGAGACAGACCCCGCCCTGAGTTCTTG 2607
Db      2880 GACATCCAGATGGCCATTGACCTTGGCGAGAGAGAGACAGACCCCGCCCTGAGTTCTTG 2939
Qy      2608 GCCAGAGCCGCCCTGACCAAGTCCCTGATGAAACCCCGGGGTGACCGAGATCA 2667
Db      2940 GCCAGAGCCGCCCTGACCAAGTCCCTGATGAAACCCCGGGGTGACCGAGATCA 2999
Qy      2668 GAGAGATGACCTGAGCGCCCTGCGGACATCCCTGAGAGCGGAGCCCGGACCGGAGAC 2727
Db      3000 GAGAGATGACCTGAGCGCCCTGCGGACATCCCTGAGAGCGGAGCCCGGACCGGAGAC 3059
Qy      2728 ACAATGTCAGCTGTGTGTCGACCCGACACACGCTCTTCATGTGATGATTCAGATC 2787
Db      3060 ACAATGTCAGCTGTGTGTCGACCCGACACACGCTCTTCATGTGATGATTCAGATC 3119
Qy      2788 GCATGGAGAAATGATGTTGTCGAGAAACCTGAGAAATTCAGAAACAGCAACGGGTGG 2847
Db      3120 GCATGGAGAAATGATGTTGTCGAGAAACCTGAGAAATTCAGAAACAGCAACGGGTGG 3179
Qy      2848 CAGAGCTGTGGTGTGTTTCAACAACTTGTGCTTCTTCTTCAACAATCACAACAGAC 2907
Db      3180 CAGAGCTGTGGTGTGTTTCAACAACTTGTGCTTCTTCTTCAACAATCACAACAGAC 3239
Qy      2908 AATCATCCCTTGTGCGACGCTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2967
Db      3240 AATCATCCCTTGTGCGACGCTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3299
Qy      2968 GAGAATCATCAGAAAGCTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3027
Db      3300 GAGAATCATCAGAAAGCTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3359
Qy      3028 AGGCGGAGAAAGCTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3087
Db      3360 AGGCGGAGAAAGCTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3419
Qy      3088 TCTGCTCGGAGCCGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3135
Db      3420 TCTGCTCGGAGCCGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3467

```

RESULT 7
BC030329
LOCUS
DEFINITION
3902 bp mRNA linear ROD 21-OCT-2003
(chondrocyte-derived), mRNA (cDNA clone IMAGE:5376197), partial
cds:
BC030329
BC030329.1 GI:20987935
VERSION
KEYWORDS
SOURCE
ORGANISM
Mus musculus (house mouse)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 3902)
Klausner, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Altschul, S.P., Collins, F.S., Wagner, C.M., Shenmen, C.M., Schlier, G.D.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.T., Wang, J., Hale, N.K.,
Datchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Sheezy, T.E., Brownstein, M.J., Ustin, T.B., Toshlyuk, S.,
Abrams, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abrams, R.D., Mulaby, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Morley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Lu, X., Gibbs, R.A.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kettner, M., Madan, A., Madan, A., Young, A.C., Shevchenko, Y.,
Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y.,

REMARK
COMMENT
JOURNAL
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
2 (bases 1 to 3902)
Direct Submission
Submitted (07-MAY-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2550,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgc@hgrl.nih.gov
Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
Hansen, N., Ho, S.-L., Karlins, E., Kong, P., Laric, P., Legaspi, R.,
Maduro, O.L., Masello, C., Maskeri, B., Mastrian, S.D., McElroy, J.C.,
McDowell, J., Pearson, R., Stantrop, S., Thomas, P.J., Touchman, J.W.,
Tsurgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
Young, A., Zhang, L.-H. and Green, E.D.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LIML at: <http://image.llnl.gov>
Series: IRAP Plate: 65 Row: m Column: 19
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis.
Location/Qualifiers
1..3902
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N-3"
/db_xref="taxon:10090"
/clone="IMAGE:5376197"
/issue_type="Mammalian tumor, MMTV-LTR/INT3 model. 5 month
old mouse. Taken by biopsy."
/clone_11b="NCI CGAP Mam2"
/lab_host="DH10B"
/note="Vector: PCMV-SPORT6"
1..3902
/gene="Farp1"
/note="synonym: Cdeq"
/db_xref="GeneID:223254"
/db_xref="MGI:2446173"
1..2510
/gene="Farp1"
/codon_start=3
/product="Farp1 protein"
/protein_id="AA030329.1"
/db_xref="GI:20987936"
/db_xref="MGI:223254"
/db_xref="MGI:2446173"
/translation="TRPEVARRLEWYGRILHPAKOREGTYKINLAVANTGILVFGFTK
INAFMAWKRIKISFPRKXRLILRPVNSYDITLRFMAGRDFCKSPFKICVHHAF
FLUREPRKRPVLFRRSSPRFSGRTGKQVLDYKGGHKVQERGHSTHRS
LVSPFPAVSEVTPKQSPQASLRFEGTSPQGSQDQKXETRACTLEHGHPSPALP
KSPGSKADVTPVPPPEEBEGGKDIKSNPDPQPSGTSLTGSFHLSELSTNS
OGGAAPANTTLSPNLSPPDKQASPLSPINDQACPTDDEEGRRKRPPTDKAVYIA
KEVSTERTYLDLVLIAVSWPQSTVSKEDSMBEALKSLIFPNPEPLHKHFTNPLKEIE

ORIGIN

Query Match 61.1%; Score 1916.8; DB 6; Length 3902;
Beet Local Similarity 86.0%; Pred. No. 0;
Matches 2151; Conservative 0; Mismatches 342; Indels 9; Gaps 2;

642 GCTCTAGAGATTGCCCCGCTAGAGATGTATGATCCGGTTGACCCGGCCAAAGGA 701
|||
5 GGCTCCGAGAGTGCCCGTCCGCTAGAGATGTATGATCCGGTTGACCCGGCCAAAGGA 64

702 CAGGGAAGGACGAGATCAATCTGGCCGTTGCCAACACGGGAATCTGTGTTTCAAGG 761
|||
65 CAGGGAAGGACGAGATCAATCTGGCTGTGCGCAACCTGGAATCTGTGTTTCAAGG 124

762 TTTCCTAGATCAATGCTTCAACTGGGCAAGGCGGGAAGCTGAGTTCAAGAGAA 821
|||
125 CTTCAGAGAGATCAATGCTTCTTAAGGCAAGGCGGGAAGCTGAGTTCAAGAGAA 184

822 GCGCTTTCTCATAGACTCCGGCCAGATCCCAATAGTGGTACACAGATACCTTGAAT 881
|||
185 GCGCTTTCTCATAGACTCCAGCTGACGTGAATACCTCATACAGAGACCTTGAAGTT 244

882 CCTGATGGCCAGTCGGGATTTCTGCAAGTCTTCTGAAAATCTGTGTGAACATATGC 941
|||
245 CCTGATGGCCGGCGGAGCTTCTGCAAGTCTTCTGAAAATATGTGTGAGAGACCATGC 304

942 CTTCTTAGACTTTTGAAGAGCCCAAGCCCAAGCCCGCTCTTTAGCCGGGG 1001
|||
305 CTTCCTCAGACTCTTGAGAGAACTTAAGCCAAAGCCCTGTTCTCTTCAAGCCGAGG 364

misc_feature
2175..2465
/db_xref="CD:smart00233"
/gene="Farp1"
/note="PH; Region: Pleckstrin homology domain. Domain commonly found in eukaryotic signaling proteins. The domain family possesses multiple functions including the abilities to bind inositol phosphates, and various proteins. PH domains have been found to possess inserted domains (such as in PLC gamma, syntenophins) and to be inserted within other domains. Mutations in Brutons tyrosine kinase (Btk) within its PH domain cause X-linked agammaglobulinemia (XLA) in patients. Point mutations cluster into the positively charged end of the molecule around the predicted binding site for phosphatidylinositol lipids"
/db_xref="CD:smart00233"

misc_feature
1695..1946
/db_xref="CD:smart00325"
/gene="Farp1"
/note="RhoGEP; Region: Guanine nucleotide exchange factor for Rho/Rac/Cdc42-like GTPases"
/db_xref="CD:smart00325"

misc_feature
1556..1556
/gene="Farp1"
/note="RhoGEP; Region: Guanine nucleotide exchange factor for Rho/Rac/Cdc42-like GTPases"

ORIGIN

Query Match 61.1%; Score 1916.8; DB 6; Length 3902;
Beet Local Similarity 86.0%; Pred. No. 0;
Matches 2151; Conservative 0; Mismatches 342; Indels 9; Gaps 2;

642 GCTCTAGAGATTGCCCCGCTAGAGATGTATGATCCGGTTGACCCGGCCAAAGGA 701
|||
5 GGCTCCGAGAGTGCCCGTCCGCTAGAGATGTATGATCCGGTTGACCCGGCCAAAGGA 64

702 CAGGGAAGGACGAGATCAATCTGGCCGTTGCCAACACGGGAATCTGTGTTTCAAGG 761
|||
65 CAGGGAAGGACGAGATCAATCTGGCTGTGCGCAACCTGGAATCTGTGTTTCAAGG 124

762 TTTCCTAGATCAATGCTTCAACTGGGCAAGGCGGGAAGCTGAGTTCAAGAGAA 821
|||
125 CTTCAGAGAGATCAATGCTTCTTAAGGCAAGGCGGGAAGCTGAGTTCAAGAGAA 184

822 GCGCTTTCTCATAGACTCCGGCCAGATCCCAATAGTGGTACACAGATACCTTGAAT 881
|||
185 GCGCTTTCTCATAGACTCCAGCTGACGTGAATACCTCATACAGAGACCTTGAAGTT 244

882 CCTGATGGCCAGTCGGGATTTCTGCAAGTCTTCTGAAAATCTGTGTGAACATATGC 941
|||
245 CCTGATGGCCGGCGGAGCTTCTGCAAGTCTTCTGAAAATATGTGTGAGAGACCATGC 304

942 CTTCTTAGACTTTTGAAGAGCCCAAGCCCAAGCCCGCTCTTTAGCCGGGG 1001
|||
305 CTTCCTCAGACTCTTGAGAGAACTTAAGCCAAAGCCCTGTTCTCTTCAAGCCGAGG 364

misc_feature
2175..2465
/db_xref="CD:smart00233"
/gene="Farp1"
/note="PH; Region: Pleckstrin homology domain. Domain commonly found in eukaryotic signaling proteins. The domain family possesses multiple functions including the abilities to bind inositol phosphates, and various proteins. PH domains have been found to possess inserted domains (such as in PLC gamma, syntenophins) and to be inserted within other domains. Mutations in Brutons tyrosine kinase (Btk) within its PH domain cause X-linked agammaglobulinemia (XLA) in patients. Point mutations cluster into the positively charged end of the molecule around the predicted binding site for phosphatidylinositol lipids"
/db_xref="CD:smart00233"

misc_feature
1695..1946
/db_xref="CD:smart00325"
/gene="Farp1"
/note="RhoGEP; Region: Guanine nucleotide exchange factor for Rho/Rac/Cdc42-like GTPases"
/db_xref="CD:smart00325"

misc_feature
1556..1556
/gene="Farp1"
/note="RhoGEP; Region: Guanine nucleotide exchange factor for Rho/Rac/Cdc42-like GTPases"

1002 GTCATCATTTCCGTTCACTGATGCTGAGTCAAGAGCAGGTTCTCGACTATGTTAAAGAG 1061
|||
365 GTTCGCTTCCTTCGCTTTAGTGGTCCGACTCAGAAAGCAAGTTCTGATATGTTAAAGAG 424

1062 AGGACATTAAGAAAGTGCAGTTTGAAGAGACAGCAAGATTTCAATCTATCCGAGCT 1121
|||
425 CCGGCAACAAGAGGTCAGTTTGAAGAAAGACAGCAAGATTCCTCCACAGAGAGCT 484

1122 TGCTTCACAGCCTTACAGAACTGAAATTCGGAAGTGTGAGAGCAGTCTCAGACAGACCG 1181
|||
485 GGTCTCAGAGCCACAGACAGCAAGATTCAGAAAGTGTGCAAAAGTCTTCCACAGAGCGCA 544

1182 CCTTACATTTGAGAAAGTGTCCGAATCTCCAGGGGCGCAAGCTCCGGCGAGAAAGGA 1241
|||
545 CCTTACATTTGAGAAAGTGTCCGAATCTCCGGGTGGCCAGAGCTCCAGCAAGCAAGGA 604

1242 ACCGAAGTTTCCGGCGGGGAGCCGGGTTCCGACCCGAGCCCTGTGGCCGAGAGAAAGCC 1301
|||
605 AACAAAGCCTGTACCTTTGAGTTGGACCTTCACCAAGCCCGGCTCTCTTAAGAGCTC 664

1302 CGCGGTTAAACAAGCGAGCGGAGCCGCTCGCGCCACAGAGAAAGAGAGAGT 1361
|||
665 TCCAGAGAGCAAGAGCTGCAAGATGAAACACGCTGTGCTGCGCTGAGAAAGAGAGAG 724

1362 CGTT-----AAGATTAGGAGCCAGCAGAGTAAACCTTACGCCCGCAGCCAAAGCAG 1415
|||
725 AAGAGGTGCAAGAGAGCGGACTCCGCGCAAGTAACTTCAGCCCGCGAGCCAAAGCAG 784

1416 CTCCCTGAGTGGACATCTTCACTTTCCGAGTGTCTGTGAATCTGCGAGGGGAGTGGC 1475
|||
785 CTCCCTGAGTGGTACCTTCACTTTCAAGAGTGTCTGTGAATCTGCGAGGGGAGTGGC 844

1476 CCTGCGCAAGCGACTTGTCTCCCACTGAGCCCGCAGCAAGAGAGGCTCTCCCT 1535
|||
845 CCGAGCAAGCGAGAGCTGTCTCCCACTGAGTCTGCAAGAGAGAGGCTCTCCCT 904

1536 GATCAGCCCGCTGCTGAATGACAGAGCGCTGCGCGGAGCGAGATGAGATGAGAGGCG 1595
|||
905 GATCAGCCCGCTGCTGAAGAGAGAGCGTGGCCAGAGCGAGAGAGAGAGGCGCG 964

1596 GAGAGAGATTTCCCACTGATTAAGCGTACTTCACTTGAAGAGAGTCTACACAGCA 1655
|||
965 GCGAAAGAGATTTCCCACTGATTAAGCGTACTTCACTTGAAGAGAGTCTACACAGCA 1024

1656 GCGAAGATATCTGAAGAGTCTGAAAGTATCACTTGTGTTTCAAGAGAGAGTACCA 1715
|||
1025 GCGAAGATATCTGAAGAGTCTGAAAGTATCACTTGTGTTTCAAGAGAGAGTACCA 1084

1716 AGAGAGCGCATGCGGAGAGCACTGAAGATCTCATATTCCGAATTTGAACCTTTGCA 1775
|||
1085 AGAGAGCGCATGCGGAGAGCACTGAAGATCTCATATTCCGAATTTGAACCTTTGCA 1144

1776 CAAATTTCAATTTCTCAAGAAATGAGAGAGAGTGTGCGCTGTGGAGAGCGCG 1835
|||
1145 CAAATTTCAATTTCTCAAGAAATGAGAGAGAGTGTGCGCTGTGGAGAGCGCG 1204

1836 CTCAATGCCCCAATCA---GAGATTACCAAAAGATTCGCGAGTCTATGCTGAAGACAT 1892
|||
1205 CTCAATGCCCCAATGAGAGAGATTCACAAAGATTCGCGAGTCTATGCTGAAGACAT 1264

1893 TCAGGGCATGAAGCACTGGGGGCTCACCTGTGGAAGCAAGCGAGGCTTGAAGGCTT 1952
|||
1265 TCAGGGCATGAAGCACTGGGGGCTCACCTGTGGAAGCAAGCGAGGCTTGAAGGCTT 1324

1953 GGAGAAATGAAGATCAAGAGCTCCGGCGGCTGGAAGACTTTCGCAAGACTTGAAGTGA 2012
|||
1325 GGAGAAATGAAGATCAAGAGCTCCGGCGGCTGGAAGACTTTCGCAAGACTTGAAGTGA 1384

2013 GAAGGTGTGTTTACCTTACCGGCTCAACACTTCTCTGCGGAGCACTGACCGGCTATGCA 2072
|||
1385 GAAGGTGTGTTTACCTTACCGGCTCAACACTTCTCTGCGGAGCACTGACCGGCTATGCA 1444

2073 CTACAAGAGAGTCTTGAGAGCGGCTGTGCAAAACACACCGCGGAGCCAGCAGCTTACG 2132

1445 CTATTAAGCATGTCCTGGAGAGGCTGTGTCAGACACACCCCAACCAACCGCGACTTCAG 1504
2133 GGACTCCGAGCCGCTTTGGCAGAGATCAAGAGATGTGTGGACAGCTCCAGCGTACGAT 2192
1505 GGACTCGAGAGCTGCGCTGGCGGATCACAGAGATGTGTGGACAGCTCCAGCGTACGAT 1564
2193 GATCAAGATGAGAAATTTCCAGAAAGCTGACGAACTCAAGAAAGATTTGATTGGCATTTGA 2252
1565 GATCAAGATGAGAAATTTCCAGAAAGCTGACGAACTCAAGAAAGATTTGATTGGCATTTGA 1624
2253 CAATCTTGTGTTCCGGGAAAGGAGTTCAATCCGTCTGGGAGGCTCAGCAAGCTCTCGGG 2312
1625 CAATCTTGTGTTCCGGGAAAGGAGTTCAATCCGTCTGGGAGGCTCAGCAAGCTCTCGGG 1684
2313 GAAGGGGCTTCAGACAGGCGATGTTCTTCTGTTTCAACGAGAGTCTGTTATACAGAGCG 2372
1685 GAAGGGGCTTCAGACAGGCGATGTTCTTCTGTTTCAACGAGAGTCTGTTATACAGAGCG 1744
2373 GGGGCTGACGGCTCCCAATCAATGTTAAAGTCAAGGGGAGCTCCGCTCATGGCATGAC 2432
1745 GGGGCTGACGGCTCCCAATCAATGTTAAAGTCAAGGGGAGCTCCGCTCATGGCATGAC 1804
2433 GATTTGAGAGAGCGAAAGACAGAGTGGGGGAGTCCCACTGCTGACCTCCGCGGCGACG 2492
1805 GATTTGAGAGAGCGAAAGACAGAGTGGGGGAGTCCCACTGCTGACCTCCGCGGCGACG 1864
2493 GCAGTCAATCATGTGTGGCGCCGAGTTCTCGGTCCGAGATGAGAAAGTGGGTTGAGACAT 2552
1865 GCAGTCAATCATGTGTGGCGCCGAGTTCTCGGTCCGAGATGAGAAAGTGGGTTGAGACAT 1924
2553 CCAGATGGGCAATGACCTGGCGGAGAGACAGAGCGCCGCTGAGTCTCTGGCGAG 2612
1925 CCAGATGGGCAATGACCTGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1984
2613 CAGCCCCCTGACAAAGTCCCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2672
1985 CAGCCCCCTGACAAAGTCCCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2044
2673 TGAATGAGAGCGCTCCGCGACATCGCTGAGAGCGCCGAGCGGACCGCGGCAACAT 2732
2045 TGAATGAGAGCGCTCCGCGACATCGCTGAGAGCGCCGAGCGGACCGCGGCAACAT 2104
2733 GAGGACGATGTGTGGACCGGAAACAGAGCTCCATGATGATGAGATTTGAGATTCGAGAT 2792
2105 GAGGACGATGTGTGGACCGGAAACAGAGCTCCATGATGATGAGATTTGAGATTCGAGAT 2164
2793 GAGGAAATCAATGTGTGGAAACCTGCTGAGAAATTTCAAAAACAGCAACGAGTGGCAGAA 2852
2165 GAGGAAATCAATGTGTGGAAACCTGCTGAGAAATTTCAAAAACAGCAACGAGTGGCAGAA 2224
2853 GCTGTGGGTGTGTTCACAAATCTTGCCTGTTCTTACAAATTCACACGAGATCA 2912
2225 GCTGTGGGTGTGTTCACAAATCTTGCCTGTTCTTACAAATTCACACGAGATCA 2284
2913 TCCCTTGGCAGGCTGCTGCTGGGCTACAGATCCGCTGAGTCCGAGAA 2972
2285 CCCCCTTGGCAGGCTGCTGCTGGGCTACAGATCCGCTGAGTCCGAGAA 2344
2973 CATTCAGAAAGATAGTGTTCAGAAAGTGTGATGAGAAAGTGTGATGAGAAAGTGTGAT 3032
2345 CATTCAGAAAGATAGTGTTCAGAAAGTGTGATGAGAAAGTGTGATGAGAAAGTGTGAT 2404
3033 GGAAGAGGATAGTGTTCAGAAAGTGTGATGAGAAAGTGTGATGAGAAAGTGTGAT 3092
2405 TGAAGAGATAGTGTTCAGAAAGTGTGATGAGAAAGTGTGATGAGAAAGTGTGAT 2464
3093 CTGCGAGACCCAGTGTGAGCAACAGAGTCTCTTGTGTA 3134
2465 CTCTCGGGCCCAATCTTGAAGTCAAAAGATCTCACTGTA 2506

RESULT 8

CO724059
LOCUS CO724059 2937 bp DNA linear PAT 03-FEB-2004
DEFINITION Sequence 9993 from Patent WO02068579.
ACCESSION CO724059
VERSION CO724059.1 GI:42284916
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
REFERENCE
AUTHORS Venter,C.J., Adams,M.C., Li,P.W., and Myers,B.W.
TITLE Kits, such as nucleic acid arrays, comprising a majority of
humanexons or transcripts, for detecting expression and other uses
thereof
JOURNAL Patent: WO 02068579-A 9993 06-SEP-2002;
PE Corporation (NY) (US)
FEATURES
source 1..2937
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 59.1%; Score 1851.8; DB 2; Length 2937;
Best Local Similarity 79.9%; Pred. No. 0;
Matches 2513; Conservative 0; Mismatches 77; Indels 555; Gaps 6;
QY 172 CAAGAGCTCTGGAGAGTGTCTGATGATGACAGTTGGCAACCACTCAACTGTGAA 231
DB 1 CAAGAGCTCTGGAGAGTGTCTGATGATGACAGTTGGCAACCACTCAACTGTGAA 60
QY 232 GGTGACTATTTTGGCTCGAGTTTCTGATACAAAAAGATCAAGGTGTGCTGATCTC 291
DB 61 GGTGACTATTTTGGCTCGAGTTTCTGATACAAAAAGATCAAGGTGTGCTGATCTC 105
QY 292 CTAAACCCATTTGTAAACAGATTAGAGCCCAACAGAGTTGTTGATGATTTGAGT 351
DB 106 -----GCCAGTTCCCGGTG 120
QY 352 AATTTCTTCGCTGACCAACATCTCAAGAAAGATCAAGATACCTGTTCGCG 411
DB 121 ACCTCTGCGCTTGTGCTGACGCCAGCTGTGATCTTCGAGATACCTGTTCGCG 180
QY 412 CTGACGTTGAGAGAGATTTGCTCAAGGAGTTGATGATGATGATGATGATGATGATGAT 471
DB 181 CTGACGTTGAGAGAGATTTGCTCAAGGAGTTGATGATGATGATGATGATGATGATGAT 240
QY 472 CTCTGATTTCAACATTTGCAATGATGATGATGATGATGATGATGATGATGATGATGAT 531
DB 241 CTCTGATTTCAACATTTGCAATGATGATGATGATGATGATGATGATGATGATGATGAT 300
QY 532 GAGCACTTGAAGAAATTAATATACATCTCAGCAAGACGCACTAGAGCAAAATCGTG 591
DB 301 GAGCACTTGAAGAAATTAATATACATCTCAGCAAGACGCACTAGAGCAAAATCGTG 360
QY 592 GAATTTACCATPACACATTTGACAAACACAGCAAGATCAGATTTCCAGCTCTAGAG 651
DB 361 GAATTTACCATPACACATTTGACAAACACAGCAAGATCAGATTTCCAGCTCTAGAG 420
QY 652 ATTGCGCGGCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 711
DB 421 ATTGCGCGGCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
QY 712 ACGAAGATCAATCTGCGCTTGCACACAGGAAATTTAGTGTTCAGAGTTTCACTAAG 771
DB 481 ACGAAGATCAATCTGCGCTTGCACACAGGAAATTTAGTGTTCAGAGTTTCACTAAG 540
QY 772 ATCAATGCTTTCACTGAGGCGCAAGGTGCGGAAGCTGAGCTTCAAGAGAAAGCTTTCTC 831
DB 541 ATCAATGCTTTCACTGAGGCGCAAGGTGCGGAAGCTGAGCTTCAAGAGAAAGCTTTCTC 600

Oy		892	AATCAAGCTCCGGGCGAATGCCAAATAAGTGTCACAGGAATCCTTGSATTCTGTATGCGGC	891
Dd		601	ATCAAAGCTCCGGGCGAATGCCAAATAAGTGTCACAGGAATCCTTGSATTCTGTATGCGGC	660
Oy		892	AGTCGGGATTTCTGCAAGTCTTCTGGAATAATCTGTGTGAACATCATGCTTCTTTAGA	951
Dd		661	AGTCGGGATTTCTGCAAGTCTTCTGGAATAATCTGTGTGAACATCATGCTTCTTTAGA	720
Oy		952	CTTTTTGAAAGCCCCMAAACCCCAGCCCCCTCTTTTAGCCCCGGGGTCAATTT	1011
Dd		721	CTTTTTGAAAGCCCCMAAACCCCAGCCCCCTCTTTTAGCCCCGGGGTCAATTT	780
Oy		1012	CGGTTCAAGTGTCGCACTCAGAAAGAGTTCTCGACTATGTTTAAAGAAGAGACAATAAG	1071
Dd		781	CGGTTCAAGTGTCGCACTCAGAAAGAGTTCTCGACTATGTTTAAAGAAGAGACAATAAG	840
Oy		1072	AAGGTGCAAGTTGTAAGAAAGACACACAAAGATTCAATTCTCCGAGCTTGTTCACAG	1131
Dd		841	AAGGTGCAAGTTGTAAGAAAGACACACAAAGATTCAATTCTCCGAGCTTGTTCACAG	900
Oy		1132	CCTACAGAACTGAATTCGGAAATGCTGGAGACAGTCTCAGCACAGACACAGGCTTACATTT	1191
Dd		901	CCTACAGAACTGAATTCGGAAATGCTGGAGACAGTCTCAGCACAGACACAGGCTTACATTT	960
Oy		1192	GGAAGAGGTGCCGAATCTTCCAAGGGGGCCAGAGCTTCGGCGAGAAAGAAAACGAAAGTTT	1251
Dd		961	GGAAGAGGTGCCGAATCTTCCAAGGGGGCCAGAGCTTCGGCGAGAAAGAAAACGAAAGTTT	1020
Oy		1252	TCCGCGCGGGGAGACC GG GGTTCGACCCGAGACCTCTGGCGGAGAGAAAGCCCCGGGTATAC	1311
Dd		1021	TCCGCGCGGGGAGACC GG GGTTCGACCCGAGACCTCTGGCGGAGAGAAAGCCCCGGGTATAC	1080
Oy		1312	AAGCAGGCGGAGCGAGACCGCCTTCGGCGGCCACGAGAGAAAGAGAGAGTGTTTAAGGAT	1371
Dd		1081	AAGCAGGCGGAGCGAGACCGCCTTCGGCGGCCACGAGAGAAAGAGAGAGTGTTTAAGGAT	1140
Oy		1372	AGGACCCAGACAGATTAACCTTAGCCCCCGACGCCAAGACAC-----	1413
Dd		1141	AGGACCCAGACAGATTAACCTTAGCCCCCGACGCCAAGACAC-----	1200
Oy		1414	-----	1413
Dd		1201	GCCAGAGGACAGATTAATTCCTTTCTTATGACTGACGTAGCTAGATAGCGAATAC	1260
Oy		1414	-----	1413
Dd		1261	GACCTTCTTAGAGAAACAAGCGTCCCGATCCCGATCCCGATCCCGATCCCGATCCCAACA	1320
Oy		1414	-----GCGTCCCTGACTGGACGTCTCTCATCTTTCCGACSTGTGTGAC	1458
Dd		1321	AATGACAAATTATGAGGGCTCCCTGACTGGACGTCTCTCATCTTTCCGACSTGTGTGAC	1380
Oy		1459	TGCGAGGGGGAGTGGCCCTTGCCCAACGTGACCTTGTCTTCCCAACTTAGGCCCGCACCC	1518
Dd		1381	TGCGAGGGGGAGTGGCCCTTGCCCAACGTGACCTTGTCTTCCCAACTTAGGCCCGCACCC	1440
Oy		1519	AAGCAGGCTCTTCCCTTGTATCAGCCCGCTGCTGTAATGACCAAGCCTGCCCCCGACGAC	1578
Dd		1441	AAGCAGGCTCTTCCCTTGTATCAGCCCGCTGCTGTAATGACCAAGCCTGCCCCCGACGAC	1500
Oy		1579	GATGAGAGTAGAGGGCCGAGAGAAAGATTTCCCACTGATTAAGCTATCTTATAGCTAAG	1638
Dd		1501	GATGAGAGTAGAGGGCCGAGAGAAAGATTTCCCACTGATTAAGCTATCTTATAGCTAAG	1560
Oy		1639	GAA GTGTCTACACCGAGCGAACATATCTGAAGGATCTCGAAGTTATCACTTCTGTGTTT	1698
Dd		1561	GAA GTGTCTACACCGAGCGAACATATCTGAAGGATCTCGAAGTTATCACTTCTGTGTTT	1620
Oy		1699	CAGAGACACGTAGAGAAAGAGAGCGCATCCCGGAAGACTGAAAGTCTATATTTCCG	1758
Dd		1621	CAGAGACACGTAGAGAAAGAGAGCGCATCCCGGAAGACTGAAAGTCTATATTTCCG	1680
Oy		1759	AATTTTGAACCTTTCACAAATTTTATATTAATTTTCTCAAGAAATTAGCAACGACTT	1818

Db	1661	AAATTTCGAACCTTTGCAACAAATTTCAATCTAATTTTCTCAAGGAAATTGACCAACGACTT	1740
Qy	1819	GCCCTGTGGAGAGGCGCCTCAATATGCCAAATCAGAGATTACAAAGAATCGCGATGTC	1878
Db	1741	GCCCTGTGCGTTGAGCTGACA-----	1761
Qy	1879	ATGCTGAAGAATTCAGGGCATBAAGACCTGGCGGCTCACTGTGTGAAGCAACAGGAG	1938
Db	1762	-----	1761
Qy	1939	GCCTTGGAGGCGCTGAGAAATGAAATCAAGAGCTCCGCGCGCTGAGAACTTCTGAGA	1998
Db	1762	-----GCTGCTAAGGGAAGAACCTTCGCTG	1786
Qy	1999	GACTTTGAGCTGCAAGAAAGTGTGTACCTACCGCTCAACACCTTCCTCTCGCGGCACTG	2058
Db	1787	CCTTCCCGCTGACAGACTGTGATGTCTTTC-----	1817
Qy	2059	CACCGGCTCATGCACTACAAAGCAGTCTGTGAGCGGCTGTGCAAAACACACCCCGGAGC	2118
Db	1818	-----	1817
Qy	2119	CACGCCGACTTCAGGAGCTGCGAGCGCGCTTTGGCAAGATCATCGAGATGTGTGCAAG	2178
Db	1818	-----CTTCAACGCGCGCTTTGGCAAGATCAAGAGATGTGTGCAAG	1860
Qy	2179	CTCCACGGTACGATGATCAAGATGAGAGAAATTTCAGAAGCTGCAGAACTCAAGAAAGAT	2238
Db	1861	CTCCACGGTACGATGATCAAGATGAGAAATTTCCAGAAGCTGCAGAACTCAAGAAAGAT	1920
Qy	2239	TTGATTGGCATTTGCAATCTTGTGTTCCTGGGAAAGGAGTTGATCGGTCTTGGGACCTC	2298
Db	1921	TTGATTGGCATTTGCAATCTTGTGTTCCTGGGAAAGGAGTTGATCGGTCTTGGGACCTC	1980
Qy	2299	AGCAAGCTCTTCGGGGAAGGGGCTCCACAGCGCATGTCTTCCG-----	2343
Db	1981	AGCAAGCTCTTCGGGGAAGGGGCTCCACAGCGCATGTCTTCCGCTGTGATGGAAGAGAC	2040
Qy	2344	-----	2343
Db	2041	GACTTGTCTCAAGGATTGTGTCACTGGGGCAAGAGAGCTCCAAAGCGGAGCTCG	2100
Qy	2344	-----TTCAACGACGTCTGTCTATACAGAGCGGGGGGCTGACGGCTCCAAATCAG	2394
Db	2101	GCAACCCAAATTCACGACGTCTGTCTATACAGAGCGGGGGGCTGACGGCTCCAAATCAG	2160
Qy	2395	TTTAAAGTCCACGGGGCAGCTCCGCGCTATGTGGCATGCAATGTGAGAGAGCGAAGACAG	2454
Db	2161	TTTAAAGTCCACGGGGCAGCTCCGCGCTATGTGGCATGCAATGTGAGAGAGCGAAGACAG	2220
Qy	2455	TGGGGGGGTGCCCACTGCTGACCCCTCCGGGGGCAAGCGGCAGTCCATCATCTGTGGCGCC	2514
Db	2221	TGGGGGGGTGCCCACTGCTGACCCCTCCGGGGGCAAGCGGCAGTCCATCATCTGTGGCGCC	2280
Qy	2515	AGTTCGTGCTCCGAGATGAGAAATGTGGTTGAGGACATCCAGATGGGCAATTGAACCTGAGC	2574
Db	2281	AGTTCGTGCTCCGAGATGAGAAATGTGGTTGAGGACATCCAGATGGGCAATTGAACCTGAGC	2340
Qy	2575	GAGAAAGACAGACCCCGCGCCCTGAGTCTCTGCGCAGACAGCCCCCTGCAACAAGTCC	2634
Db	2341	GAGAAAGACAGACCCCGCGCCCTGAGTCTCTGCGCAGACAGCCCCCTGCAACAAGTCC	2400
Qy	2635	CCTGATGAAGCCACCGGGGCTGACACAGAGTCAAGAGATGACTTGAGACCGCTCGCGACA	2694
Db	2401	CCTGATGAAGCCACCGGGGCTGACACAGAGTCAAGAGATGACTTGAGACCGCTCGCGACA	2460
Qy	2695	TGCGTGAAGCCACAGGCCCGCAACCGGGGCAACAATGTGTGCACTGTGTGTGCAACGC	2754
Db	2461	TGCGTGAAGCCACAGGCCCGCAACCGGGGCAACAATGTGTGCACTGTGTGTGCAACGC	2520
Qy	2755	AACACAGCGTCTCCATGTGTGAATTGACA-----TCGAGTGAAG	2796

proteins. PH domains have been found to possess inserted domains (such as in PLC gamma, syncrophins) and to be inserted within other domains. Mutations in Brutons tyrosine kinase (Btk) within its PH domain cause X-linked agammaglobulinemia (XLA) in patients. Point mutations cluster into the positively charged end of the molecule around the predicted binding site for phosphatidylinositol 1phd8"

/db_xref="CDD:smat00233"

ORIGIN

Query Match 33.3%; Score 1043; DB 6; Length 2632;

Best Local Similarity 89.3%; Pred. No. 2.6e-282;

Matches 1124; Conservative 0; Mismatches 135; Indels 0; Gaps 0;

```
QY 1876 GTGATGCTGAAGAACTTTCAGGGCATGAAGCACTTGGCGCTCACTGTGGAAAGCAAGC 1935
DB 1 GTCATGCTGTAATAAACTTCAGGGCATGAAGCACTTGGCGCTCACTGTGGAAAGCAAGT 60
QY 1936 GAGGCTTGGAGGCTCTGGAGATGGATCAAGAGCTCCGGCGGCTGGAGAACTTCTGC 1995
DB 61 GAGGCTTGGAGGCTCTGGAGATGGATCAAGAGCTCCGGCGGCTGGAGAACTTCTTGC 120
QY 1996 AGAGACTTTCAGCTGCAAGAGGTGTCTTACCTACGCTCAACACTTCTCTGCGGCCA 2055
DB 121 CGAGACTTTCAGCTGCAAGAGGTGTCTTACCTACGCTCAACACTTCTCTGCGGCCG 180
QY 2056 CTGCACCGGCTCATGATCACTCAAGAGGTCTTGGAGGCTGTGCAAAACCAACCGGCCG 2115
DB 181 CTGCACCGGCTCATGATCACTCAAGAGGTCTTGGAGGCTGTGCAAAACCAACCGGCCA 240
QY 2116 AGCCAGCGGACTTTCAGGAGCTGCGAGCGCTTTGGGAGATGCAAGAGATGGTGGCA 2175
DB 241 AACCAAGCGGACTTTCAGGAGCTGCGAGCGCTTTGGGAGATGCAAGAGATGGTGGCC 300
QY 2176 CAGCTTCACGATGATGATCAAGATGAGAAATTTCCAGAGCTGACGAACTCAAGAA 2235
DB 301 CAGCTTCACGATGATGATCAAGATGAGAAATTTCCAGAGCTGACGAACTCAAGAA 360
QY 2236 GATTTGATTTGGATTTGATGAAATTTTGTGCTTCCGGAGAGGATTTATCCGCTTGGGAGC 2295
DB 361 GATTTGATTTGGATTTGATGAAATTTTGTGATCCAGAGAGGAGATTTATCCGCTTGGGAGC 420
QY 2296 CTACAGCAAGCTCTCGGGAGAGGGGCTCAAGAGCGCATTTCTTCTGTTCAACGACGTC 2355
DB 421 CTACAGCAAGCTCTCGGGAGAGGGGCTTCAAGAGCGCATTTCTTCTGTTCAACGATGTC 480
QY 2356 CTGCTATACAGAGCCGGGGGCTGACGCGCTTCCATCACTTTAAAGTCCACGCGACGCTC 2415
DB 481 TTGCTGTATACAGAGCCGGGGGCTGACGCGCATTTCAATCACTTTAAAGTCCACGCGACGCTC 540
QY 2416 CGGCTCTATGGATGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2475
DB 541 CAGCTCTATGGATGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
QY 2476 ACCCTCCGGGGGCGAGCGGAGTTCATCATGTCGCGCGCAGTTCTCGGTCCGAGATGAG 2535
DB 601 ACCCTCGAGGGGCGAGCGGAGTTCATCATGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 660
QY 2536 AAGTGGGTTGAGAGATCCAGATGGCATTTGACCTTGGCGAGAGAGAGAGAGAGAGAGAG 2595
DB 661 AAGTGGGTTGAGAGATCCAGATGGCATTTGACCTTGGCGAGAGAGAGAGAGAGAGAGAG 720
QY 2596 CCGTAGTTTCTGGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2655
DB 721 CCGTAGTTTCTGGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
QY 2656 GACCAAGAGTCAAGAGATGATGAGCGCTGCGGCACTTCGCTGGAGGCGGAGCGCCCG 2715
DB 781 GACCAAGAGTCAAGAGATGATGAGCGCTGCGGCACTTCGCTGGAGGCGGAGCGCCCGT 840
QY 2716 CACCGCGGCAACAAATGTTGACGCTGTGCTGGCAACCGCAACCAACGCTTCTTCATGTTG 2775
```

```
DB 841 CACCGTGCAACACGATGATGACGTTGCTGGCATTCGACAGCAACGATGTCATGATG 900
QY 2776 GACCTTACCATTCGATGATGAGAAATTCAGTTGCTGGAACCTGCTGAGAAATTCAGAAAC 2835
DB 901 GACCTTACCATTCGATGATGAGAAATTCAGTTGCTGGAACCTGCTGAGAAATTCAGAAAC 960
QY 2836 AGCAACGGGTGGCAGAGCTGTTGGGTGTTTCAAACTTTCGCTGTTCTTTCACAA 2895
DB 961 AGCAATGGCTGGCAGAGCTGTTGGGTGTTTCAAACTTTCGCTGTTCTTTCACAA 1020
QY 2896 TCACACAGAGATCATTCATCCCTTGGCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2955
DB 1021 TCACACAGAGATCATTCATCCCTTGGCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
QY 2956 CCGCTGAGTCCGAGAGATCAGAAAGATTCAGGTTCAAGCTGACCTTCAAGTCCGAC 3015
DB 1081 CCGCTGAGTCCGAGAGATCAGAAAGATTCAGGTTCAAGCTGACCTTCAAGTCCGAC 1140
QY 3016 GTCTACTTCTCAGGCGGAGAAAGCGATGACAGTTGCAAGAGTGGATGGAATGATCCG 3075
DB 1141 GTCTACTTCTCAGGCGGAGAAAGCGATGACAGTTGCAAGAGTGGATGGAATGATCCG 1200
QY 3076 AGTGCACCAAGCTGCTGCTGCGGACCCGACGCTGTTGAGCCACAAAGATCTTGTGTA 3134
DB 1201 AGTGCACCAAGCTGCTGCTGCGGACCCGACGCTGTTGAGTGCACAAAGATCTTCA 1259
```

RESULT 10
LOCUS CQ725044 3997 bp DNA linear PAT 03-FEB-2004
DEFINITION Sequence 10978 from Patent WO02068579.
ACCESSION CQ725044
VERSION CQ725044.1 GI:42285901
KEYWORDS

SOURCE

ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.

REFERENCE

AUTHORS Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
TITLE Kites, such as nucleic acid arrays, comprising a majority of
humanecons or transcripts, for detecting expression and other uses
thereof
JOURNAL Patent: WO 02068579-A 10978 06-SEP-2002;
PE Corporation (NY) (US)

FEATURES

source location/Qualifiers
1..3997
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 33.1%; Score 1039; DB 2; Length 3997;

Best Local Similarity 60.3%; Pred. No. 3.9e-281;

Matches 1865; Conservative 0; Mismatches 1185; Indels 45; Gaps 7;

```
QY 21 GCCGACCCAGAGATGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 80
DB 150 GCGAGCTGAGAGATGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 209
QY 81 TGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 140
DB 210 TGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 269
QY 141 GAGTACACCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 200
DB 270 GAGCAACACCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 329
QY 201 TGCAGTTGCAACCACTCACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 260
DB 330 ACAAGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 389
```

QY 261 TCACAAAAGATCAGCGGTGGCTGATCTCTTAAACCCATTTGTGAACGATTAGAG 320
DB 390 TACTCACTCTTACGTGATTTGGCTTGAACCTATBAAACCATCTTTAGGCAAAATACAGAG 449
QY 321 GCCAAGACAGTTGTTGTTAGTTGTGTGAATTTCTTTCCGCTGACACACAACT 380
DB 450 GCCAAGAAATGTGGTGTCTGCTAGCTGTAAATTTTTCACCTGATCCTGGTACGT 509
QY 381 CCAAGAAAGATCACAAGGTAACCTGTTCCGGCTGCAAGTGAAGAGAGACTTGGCTCAAG 440
DB 510 ACAAGAAAGATATACAGGATCTTGTGTGCTTGCACTTAAGAGACCTGCTGAGA 569
QY 441 CAGGTGACGTGTAATGACACAGCGAGCTCTTGTATTTCACACTTGTGCAATGTA 500
DB 570 GCGTTGACCTGTGCTGACACACAGCGGCTTCTCAAGTCCCATCTCTGCACTGGA 629
QY 501 GATTGGGATTTTGTATGAGCCTTGAACAGAGACCTATGACAAAATAATATCATACC 560
DB 630 AATAGGAGATTACGATGAAAGCTGACCGAGACACTCAAGTGAACGATATTTGCC 689
QY 561 TCAGCAAGACGACTAGAGACAAATCTGTGAATTTTCACTATACCACTTGGACAAAC 620
DB 690 TGGCCAGACGACTGCTTGAAGAGATATGAAATTCATCAGAAAGACCTGGGCCAGAC 749
QY 621 ACCAGCAATTCAGATTTCCAGCTCTAGAGATTGCCGCTGAGTAGATGTAATGGAAT 680
DB 750 ACCTGTGAGTCCGATTTTCCAGGTGCTCGAAATTTGCTGAAAGTTGGAAATGACGAT 809
QY 681 CCGGTTGACCCGSCCAAGACAGGAAAGCAGAGATCATCTGGCCCTTGCACAC 740
DB 810 CAGATTTCAATGCTTCTGACAGGGAAGAACCAAGATTCATGCACTTCCCAT 869
QY 741 GGGAAATCTAGTGTTCAGGGCTTCACTAAGATCAATGCTTCACTGGGCCAAGTGG 800
DB 870 GGGGTGATCTGTGTCCAGGGCACACCAAAATCAACATTCATCTGATCTCAAGGTCG 929
QY 801 GAGGTAGGCTTCAAGAGAAAGGCTTCTCATCAAGCTCCGGCAGATGCCAATATGTGC 860
DB 930 TAAACTATAGCTTCAAGAGAAAGATTTCTTATCAACTTCATCCAGGTTCAATGACC 989
QY 861 GTACAGAGATACCTTGAATCTCTGATGCGCAGTCCGGAATTTTGCAGATCTTCTGAA 920
DB 990 TTACAGAGACACTTGAATTTTGTGTGATGAGATGAAATGTAAAGAACTTCTGAA 1049
QY 921 AATCTGTGTGAACATCATGCTTCTTGTGAACCTTTTGAAGAGCCCAACCAAGCCCA 980
DB 1050 GATTTGTGTGAGATCAACCTTTTGTGAACCTTTTGAAGCAACTTAAGCCAAAGCAAA 1109
QY 981 GCCGCTCTCTTAGCGGGGGGTCACTATTCCGTTCAGTGTCCGACTCAGAGCAGGT 1040
DB 1110 AGCCGCTCTTCAAGCGGGGCTCTCTTCAATACAGTGAAGAACTCAGAAACAACT 1169
QY 1041 TCTCGACTATGTTAAAGAGAGACATTAAGAGGTGCACTTGTGAAGAGAGACACAA 1100
DB 1170 AGTAGATTTATTTCAAGAGAGGTGAATGAAGAAATTCATATGAAGAGAGACAGCAA 1229
QY 1101 GATTCATTTCTATCCGAGCTTCTTCAACGCTTACAGAACTGAATTCGGAAGTGTGGA 1160
DB 1230 GACCCACAGCTC-----GTTCCAGCTCTGACTGCAAGACCTTACCAAA 1271
QY 1161 GCAGTCTCAGACAGACAGCCTTACATTTTGAAGAGGTGCCGAATCTCAGAGGGCCA 1220
DB 1272 ACAGAGACTCTATTTCCCGAGGAGATTGAAGACTCTGCTCCCACTTTCACAGAAATGC 1331
QY 1221 GAGCTGCGCGGAGAGAAAGAACCGAAGTTTCCGCGGGAGCCGAGGTGCGACCCGAG 1280
DB 1332 CTTTTCGCTCTCTCCCTCACTGCTGCTCCCTGAGCTCCAGAGTTTAAAGACAG 1391
QY 1281 CCTTGCGCGGAGAGAGAGCCCGCGGGTAAACAGCAGGCGAGAGAGCGCTCGCGCC 1340
DB 1392 CAGCAGCTCTCTTCAAGATCTCCAGAGTTTCTAGCTTAAAGATCCAGCTGACAGAGGCG 1451
QY 1341 CACGAGAGAGAGAGAGG---TCGTTAAGATTAGGACCCAGCAGAGTAACTCAGCC 1397

DB 1452 CAGTGAAGCAGTGGCTGAGAGCCCCCAACACACATGCGGCCACGCCCTCGGGCCCCCGC 1511
QY 1398 CCGCAGCCAGACAGAGCTCCCTGACTGGCAGTCTTCACTTTCCGAGCTGTGTGAA 1457
DB 1512 ACTTCAGCCTGTGTCCAGGCTTTTCAAGAGATCTCAGCCTTCTCCCTCCAGCGGAA 1571
QY 1458 CTGCAAGGGGGAGTGGCCCTTCCAGAGTACCTGTCTCCCACTGAGCCCGACAC 1517
DB 1572 GAGCCCTGAGTCTGAGCCCTGCAATTTCAAGTCCCT-----TGGCCCACTGTA 1622
QY 1518 CAAGCAGCCTCTCCCTGATTCAGCCGCTGTGAAATGACAGGCTGCCCCGAGACGA 1577
DB 1623 ACAGGCTATCTCCCACTCTTGAAGCTTGTCTCACTGATGCTGGCGAGGCCGAGATGA 1682
QY 1578 CGATGAGATGAGAGGCGGAGAGAGATTCGCACTGATTAAGCGTACTTCACTAA 1637
DB 1683 CTGCGAG---GAGCCAGACACAGGCGCGTCTGACAGAGAGGCTTACTTCAATGCAA 1739
QY 1638 GGAAGTGTCAACACCGAGGAAATATCTGAAAGATCTGAAATTAATCACTTCGTGTT 1697
DB 1740 AGGATTTCTGCTTACAGAAAGAACTATCTCAAGATTTAAGATTAATTCGTGTGTT 1799
QY 1698 TCAGACACAGTGAAGAGAGAGAGCAGCATGCGGAAAGCACTGAAAGTCTCATATTCC 1757
DB 1800 CCGCAGCGAGTGTGAAGAGAGAGCCGATGCTGCGACTGTATGACGCTGCTTTCTC 1859
QY 1758 GAAATTTGAACCTTTGACAAATTTTCAATATTTCTCAAGAAATTGAGCAACGACT 1817
DB 1860 CAACATCGATCCCATATGATGATTCACAGAGGCTTCTGCGGAGGTGAGAGAGGCT 1919
QY 1818 TGGCTGTGGGAAGGCGCTCAAAATGCCAAATCAAG---ATTACCAAAAGATTGGGCA 1874
DB 1920 GGCACCTGGAAGAGGCGCTCCCAAGCCCAACAAAGGAGATCAACGAATCGGGGA 1979
QY 1875 TGTCTGTGGAAGAACTTCAAGGAGATGAAAGCACTGCGGCTCACCTGGAAGACAG 1934
DB 1980 CATCTGCTCAGGAACATGCGGATGAAAGAGTTTACAGCTACTTCCAAATACATGA 2039
QY 1935 CGAGGCTTGAAGAGCCCTGAGAGATGAATCAAGAGCTCCCGGCGCTGAGAACTTCTG 1994
DB 2040 CGAGGCTCTTAACGAAGTGAAGAGGCTTACCAAGCCTGTAAAGAAATGAGGCAAGTGA 2099
QY 1995 CAGAGCTTTGAGCTGCAAGAGTGTGTTACCTACCGCTCAACACTTCTCCTGGGGCC 2054
DB 2100 CAGGAGTTTGAAGCTGCAAGAGTGTGCTTACTTCTCAACAGTCTCTGTAAGCC 2159
QY 2055 ACTGCAAGGCTCATGCACTAACAGAGGCTCTGAGCGGCTGTGCAACACCAAGCCGCC 2114
DB 2160 CATCCAGCGGCTGTGCACTACCGCTGCTGTGCGGCCCTGTGTGGGACATTAAGGCC 2219
QY 2115 GAGCCAGCCCACTTCAAGGAGCTGCCAGCCGCTTTTGAAGAGATCAAGAGATGAGTGGC 2174
DB 2220 CGGGCACATGACTAGCTGATCGCATGACGCTTGAAGGATCAACAGAGGTGACAC 2279
QY 2175 ACAGCTCCAGGTAAGATGATCAAGATGAGAAATTTCCAGAGCTGCAAGAACTCAAGAA 2234
DB 2280 CACACTACAGCACTTCTATCCGCTGAGAACTGCAAGAACTTACGAGAGCTGACAGG 2339
QY 2235 AGATTTGATTTGAGATGACATCTTGTGTTCCGGGAAAGGAGTTATCCGTCTGGGAG 2294
DB 2340 GAGCTGTGTGGGATTAAGAACTTCTGCTCTGAGAGGAGTTATCCGTGAGGGCTG 2399
QY 2295 CTTGAGCAAGCTCTCGGGAAAGGGGCTCAGACAGCGCATGTTCTTCTGTTCACAGCT 2354
DB 2400 CTTTCAAGAGCTCACCAAGAGGGCTGTGACAGAGAGAGTGTCTTCTGTCTCAGATAT 2459
QY 2355 CCTGCTATACAGAGCGGGGCTGACGCGCTCCATCACTGTTAAAGTCAAGCGGAGCT 2414
DB 2460 GTTGTGTACAGAGAAAGAGTTGACAGGAGACAGGCCCTTCCGAGTCCGGGGCTCTCT 2519
QY 2415 CCGGCTCTATGCAATGAGATTTGAGAGAGCAAGAGAGTGGGGGTGCCACTGCTCT 2474

Db 2520 TCCCTCCAAAGCATCTGTGGAGAAAGATGAACAGATGTCCTTCCACTGTTT 2579
QY 2475 GACCTCTCGGGGCGACGGGCAATTCATTCGTGGCCGCAATTCCTGGTCCGAGATGA 2534
Db 2580 CACCACTCAAGCGGCTCAAGAAAACAATCGTGGGAGCAGACCTCGGCTGAGAAA 2639
QY 2535 GAAGTGGATTGAGGACATCCAGATGCGCAATTGACCTGGCGGAGAGAGAGAGACCCCGC 2594
Db 2640 GAAGTGGATGCTGGACCTGAATCTCCCGCATCCAGAGACCAAGAGTGGGGTGAACCGGC 2699
QY 2595 CCTGTGATTCCTGGCCAGCA-----GCCCTCTGCAACAAGTCCCTGATGAGCCAC 2648
Db 2700 CCTGTGACCTGCAAGGCGCCGACTGTGTGCACTCGTCCCGCCAGATCCCGCAAGAGGTATC 2759
QY 2649 CCGCGGTGACAGAGATCAGAGATGACCTGAGCGGCTCGGCGACATCGCTGAGAGGCGCA 2708
Db 2760 TCTG---GAGCAGAGATCAGAGATGATCTCGGGGTGTCCGAGCTCCCTGGAGGGGCA 2816
QY 2709 GCGCCCGCACCGCGGCAACAATGCTGACGTCGTCTGSCACCGCAACAGCAGCTCTC 2768
Db 2817 TGGCCAGCACCGGGCCACACCAATGACGTCGTCTGTGTACCGGAGACACGAGCTGTC 2876
QY 2769 CATGTGACTTTCAGATCGCAGTGGAGATCATGTTGTTGAAACCTCTGAGAAATT 2828
Db 2877 CAGGGCAGACCAAGTGCAGCTGTGAGAAACGATTCAGAGATCTGTGAGAAAGTT 2936
QY 2829 CAAAAACAGCAACGGGTGGCAAGAGCTGTGGGTGTGTTCACAAATTCTGCTGTTCTT 2888
Db 2937 CAAAAACATCATGCTGGCAGAGAGCTGTGGGTGTGTTCACAAATTCTGTTGTTCTT 2996
QY 2889 CTACAATATCACACAGAGACATATCCCTTGGCAGCCTGCTCTGCTGAGTACGCT 2948
Db 2997 CTACAAATCATATGAGATGATCACTACCTGGCAGCCTCCGCTGCTGGGCTACAGGCT 3056
QY 2949 CACCATCCCTCTGAGTCCGAGAACATCCAGAAAGCTACGTGTTCAAGCTGACTTCAA 3008
Db 3057 GAGCATCCCCAGGAGGCGGATGAGTACAGAAAGCTATGTTTCAAGCTCAGATTTCAA 3116
QY 3009 GTCCCACTCTACTACTTACAGGGCGGAAAGCGATCAGGTCGAAAGGTGATGGAAT 3068
Db 3117 ATCCCACTCTACTTCTCCGGCTGAGAGCAAGTACAACTTTGAAAGGTGATGAGGT 3176
QY 3069 GATCCGAGTGGCCACAGCTGTGCTCGGACCCC 3103
Db 3177 GATCCAGGGGGCCAGCAGCTCAGCCGGGAGGGGCC 3211

RESULT 11
CO895396 3997 bp DNA linear PAT 05-NOV-2004
LOCUS Sequence 16 from Patent EP1471153.
DEFINITION CO895396
ACCESSION CO895396
VERSION CO895396.1 GI:55467577
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE
1 Certi, U., Foser, S. and Weyer, K.
AUTHORS Transcriptional activity assay
TITLE Patent: EP 1471153-A 16 27-OCT-2004;
JOURNAL F. HOFMANN-LA ROCHE AG (CH)
FEATURES
Location/Qualifiers
1..3997
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 33.1%; Score 1037.4; DB 2; Length 3997;
Best Local Similarity 60.2%; Pred. No. 1.1e-280;

Matches 1864; Conservative 0; Mismatches 1186; Indels 45; Gaps 7;
QY 21 GCCGACCCAGGATATGAGATCTGGGGGCCCCGAAAAATTCGGGGATCAGTACCTTGGAAACG 80
Db 150 GCGACTGAGGAGATGCCCTTGGGTGCGAGACCCCTGTGGAGTTTGAACCTTGAAGCC 209
QY 81 TGGACAGAACCGCCCCCAACCTTCAGAAAACTGTGTTCATCAAAATTCAGATGCT 140
Db 210 TGGGAGACTCTCTTGCCAGAAATGCAAGAGAGACCTGCACCTCAGATGAAGCTGCT 269
QY 141 GGATGACACCCAGAGGAGCATTTGAAGTTCCACAAAGAGCTCCTGGAAAGTCTCTGTA 200
Db 270 GGACAAACACATGGAATATTTGACATTGAGCTTAAAGCGATGCCAGATTTACTGAC 329
QY 201 TGCATTTGACACACACTCAACTCGTGGAAAGTACTATTTTGGCTCGAGTTTCTGA 260
Db 330 ACAGTGTGAAAGCTTTTAACTGTGTGAATGTACTACTTTCGGAGTGAATTTCAAAA 389
QY 261 TCACAAAAAGATCAGCGTGTGCTGATCTCTTAAACCCATTTGTGAACAGATTAGAG 320
Db 390 TACTGAGTCTACTGAGATTTGGCTTGAACCTATGAACCCATCATTTAGCAAAATGAGAG 449
QY 321 GCCAAAGCAGCTTTGTTGTTAAATTGTGTGAATTTCTTTCGCTGACACACACT 380
Db 450 GCCAAAGATGTGTCTTGCTTGTGCTGATGCTTAAATTTTTCACCTGATCTGTGAGCT 509
QY 381 CCAAGAAAGATCACAAGGTACTGTGCGCTGACGGTGAAGAGACGACTTGCTCAAG 440
Db 510 ACAGAGATATTAAGATATCTTTTGTCTTGCACCTTAAAGAGACCTCTGTGAAGA 569
QY 441 CAGTTGACGTGTAAATGACACAGGAGCTCTTGTGATTTTCACATTTGCAATTGTA 500
Db 570 GCGTTTGAAGTGTGTGACACACAGCGGCCCTTCTCAGTCCCATCTCTGACATGGA 629
QY 501 GATTGGGATTTTATGAAAGCTTGTGAAGAGACCTTGAAGAAATTAATATATACC 560
Db 630 AATAGAGATTTAAGATGAAGCGCTGAGACCGAGAGACCTCAAGGTAAAGATATTTGGC 689
QY 561 TCAGCAAGCGCATAGAGAGCAAAATCGTGAATTTTCAACCTAATCACTTGGACAAAC 620
Db 690 TGGCAGACGACTGCTTGAAGAAATCTAGAAATTCATCAAGAGCAGTGGGCCAGAC 749
QY 621 ACCAGCAATATCAGATTTTCAGCTCTTGAAGATTTGCCGTGCGCTAGAGATGATGAAT 680
Db 750 ACTGCTGATGTGATTTCCAGGTGCTCGAAATTCGAAAGTTGGAATGTACGGCAT 809
QY 681 CCGGTTCACCCGGGCAAGAGACAGGAGAGCAAGAAATCAATCTGGCCGTTGCCACAC 740
Db 810 CAGATTTCAATGTGCTTCTGACAGGGAAGAAACCAAGATTCAACTGGCAGTTTCCACAT 869
QY 741 GGGATTTCAAGTGTTCAGGGTTTCACTAAGATCAATGCTTCACTGGGCGCAAGGTGCG 800
Db 870 GGGTGTACTCGTGTTCAGGGGACACCAAAATTCAACTTTCAACTGTGCTCAAGGTCCG 929
QY 801 GAGGTGAGCTTCAAGAGAAACGCTTCTCATCAAGCTCGGGCAGATGCCAATATGTC 860
Db 930 TAAATTAAGCTTCAAGAGAAAGATTTCTTATCAAACTTCACTCAGAGGTTCAATGACC 989
QY 861 GTACAGAGATCTTGGAAATTCGTATGCGCAGTGGGATTTTCGCAAGTCTTCTGGA 920
Db 990 TTACACAGACATTTAGAAATTTGTGTGGTATGAGAAATGAAATTTCAAACTTCTGGA 1049
QY 921 AATCTGTGTGAACATCATGCTTCTTATGACTTTTGAAGGCCCAACCAAGCCCA 980
Db 1050 GATTTGTGTGAGATTCACACTTTTATGACTTTTGAACCACTTAAAGCCAAAGCAAA 1109
QY 981 GCGGTCTCTTTAGCCGGGGGTCTATCATTTGCTTCAAGTGTGCACTCAGAGCAGGT 1040
Db 1110 AGCGCTTCTTCAAGCCGGGGGTCTCTCTCAAGATACAGTGAAGAACTCAGAAACAAT 1169
QY 1041 TCTGACATATGTTAAAGAGAGACATTAAGAGTGCAGTTTGAAGAGACACAGCA 1100
Db 1170 AGTAGATTATTTCAAGACAGTGAATGAGAAATTCATATGAAGAGGACACAGCA 1229

QY 1101 GATTCAATTCTATCCGAGCCTTGCTTCAACGCTTACAGAACTGAATTCCGAAGTGTGGA 1160
DB 1230 GACCCACACGCTC-----GTTCCAGCTCTGACATGCAACCTTACCAAA 1271
QY 1161 GCAGTCTCAGACAGACACCAAGCTTACATTTGGAGAGTGTCCGAAATCTCCAGGGGGCCA 1220
DB 1272 ACAGAGATCTCATTTCCCGAGGGGATTGAGGACTCTCGCTCCCACTTTCACAGAAATGC 1331
QY 1221 GAGCTGCGGGCGAGAAAGAACCGAAGTTTCCGCGGGAGCCGGGGTGCACCCGAG 1280
DB 1332 CTTTTACTGCTCTCTCCCTCCACTCTGCTGCTCCCTGCGCTCCAGAGTTTAAAGCAG 1391
QY 1281 CCTGCGCGGAGAGAAAGCCCGCGGTAAACAAGAGCGGAGCGAGCCGCTCGGCGCC 1340
DB 1392 CAGCAGCTCTCTCAAGATCCCAAGTTTCTTACGTCAAGATGTCAGCTGCAAGAGGCG 1451
QY 1341 CACGAGAGAAAGAGAGAGG---TCCGTTAAGATAGGACCAAGATTAACCTCAGCC 1397
DB 1452 CAGTGAAGCAGTGGCTGGAGGCCCCGACACACATCGGCCACGCCCCCGGGCCCCCGC 1511
QY 1398 CCGGACGCCAAGACAGGCTCCCTGACTGGCAGTCTTACCTTTCCGAGCTGTCTGAA 1457
DB 1512 ACTCCAGCTGCTCAGGCTTTTCAAGAGAGTCTTCAAGCTTCTCCCTCCAGCGGAA 1571
QY 1458 CTCGCAAGGGGGAGTGGCCCCCTTCCAACTGACCTTGTCTCCAACTGAGCCCCGAC 1517
DB 1572 GAGCCCCCTGAGCTGAGCCCTTCAATTTGAGTGCTT-----TGGGCCCACTGTA 1622
QY 1518 CAGCAGGCTCTCTCCCTTGATCAGCCGCTGTGAATGACAGAGCTTCCCGGACGGA 1577
DB 1623 ACAGGGCTATCTCCCACTCTTGAGCCCTGTCTCTCAGTGAAGTGTGGGAGGACGGATGA 1682
QY 1578 CGATGAGATGAGGGCCCGAGAGAGATTTCCCACTGATTAAGCTTCTTCACTAA 1637
DB 1683 CTGCGAG---GAGCCAGACACAAAGCGCTGCTGAGAGAGAGCTTCTTCAATGCA 1739
QY 1638 GGAAGTCTACCAACCGAGGAACTATCTGAGAGATCTGAGATTATCTCTGCTGTT 1697
DB 1740 AGGATTTCTGCTACAGAAAGAACTATCTGAGAGATTGAAAGTTTACCGTGTGTT 1799
QY 1698 TCAGACACAGTAGAGCAAAAGAGAGCGCATGCGGAAAGCACTGAAAAGTCTCATATTCC 1757
DB 1800 CCGCAGCGCAGTGTGTAAGAGAGAGCCCATGCTGCGACTGTATGACGTCTTCTTC 1859
QY 1758 GAATTTTGAACCTTTGCACAAAATTTCACTAAATTTTCTCAAGAAATTGAGCAAGACT 1817
DB 1860 CAACATCGATCCCATCTATGATTTCCACAGAGCTTCTGCGAGGTGAGCAAGGCT 1919
QY 1818 TGCCCTGTGGAAAGCCGCTCAAAATGCCAAATCAGAG---ATTACCAAGATTCGCGA 1874
DB 1920 GCGACTCTGGGAAAGGCGCTTCCAAAGCCACAAAGAGCAGTCAATCAAGATTCGGGA 1979
QY 1875 TGTCATGCTGAAGAACATTCAGGGCATGAGACCTGCGGCTCACCTGTGAGAGCAG 1934
DB 1980 CATCTGCTCAGAGACATGCGCAATTAAGAGTTTACAGCTTCTTCCAAAGACATGA 2039
QY 1935 CGAGGCTTTGAGAGCCCTGAGAGATGAATCAAGAGTCCCGCGCTGAGAACTTCTG 1994
DB 2040 CGAGGCTTAAACGAACTGGAAAGGCTAACAAACGCTGTAAGAGTTGAGGCAAGTGA 2099
QY 1995 CAGAGCTTTGAGAGTCAAGAGTGTGTTACCTACCGCTCAACCTTCTCTGCGGCG 2054
DB 2100 CAGAGTTTGAAGCTGAGAAAGTGTGTACTTGTCTTCAACAGTTCTGTCTGAAGCC 2159
QY 2055 ACTGCAACGGGCTCATGCACTACAGCAGGTCTGAGCGGCTGTGCAACCAACCGCGC 2114
DB 2160 CATCAGCGGCTCTCTGCACTACCGGCTGTGTGCGCGGCTATGCGACATTAAGCC 2219
QY 2115 GAGCAGCGCGACTTCAGGAGCTGCGAGCGCTTTGGCAGAGATCAAGAGATGTGTGC 2174
DB 2220 CCGGACCATGACTAGCTGAGCTGACAGCCCTGAAAGCATCAAGAGGTGACAC 2279

QY 2175 ACAGCTCAAGGATGATGCAAGATGAGAAATTTCCAGAACTGACAGAACTCAAGA 2234
DB 2280 CACACTACAGACATTTCTATCTCGGCTGGAGAACTGCAAACTTACGAGTGTGACGG 2239
QY 2235 AGATTTGATTTGCAATTAATCTTGTGTTCCGGAGAGGAGTATCTCTGTGGCAG 2294
DB 2340 GAGCTGTGTGGGATGAGAACTCATTTGCTCTTGGCAGGAGATTCATCGTGAAGGCTG 2399
QY 2295 CCTCAGCAAGCTTCGGGGAAGGGGCTCCAGACGCGCATGTTCTTCTGTTCACAGACGT 2354
DB 2400 CTTTCAAACTTCAACMAAGGCGCTTGCAGCAGAGATGTTTTTCTGTTCACATTA 2459
QY 2355 CCGTATTAACAGAGCGGGGCTGACGGCTCCATTCAGTTTAAAGTCCAGCGGAGCT 2414
DB 2460 GTTGTCTTACAGAAAGAGAGTTCAGAGACACACCTTCGGATTCGGGGCTTCT 2519
QY 2415 CCGCTCTATGCGATGCAATTTGAGAGAGCGAAGACAGTGGGGGTGCCCCATGCT 2474
DB 2520 TCCCTTCCAAGGATCTGTGTGGAAGAAAGTATTAAGATGTCTTCCACATGTTT 2579
QY 2475 GACCTTCGGGGCCAGCGGCACTCATCTGTGCGGCAATTTCTCGTCCGAGATGA 2534
DB 2580 CACCATCTACGCGCTCAGAAACAAATCGTGTGCGAGCGACACTCGCTGGAAGAA 2639
QY 2535 GAAAGTGTGAGGACATCCAGATGCGCACTTGAAGCTGCGGAGAAAGAGCAGCCCGC 2594
DB 2640 GAAAGTATGCTGAGACTTAATCTCGCGATCCAGAGCCAGAGAGTGTGACACGCG 2639
QY 2595 CCCTGAGTTCTTGGCCAGCA-----GCCCTCTGCAACAAAGTCTCTGATGAACAC 2648
DB 2700 CCTGCACTGCAAGCGCGGACTGTGTGCACTGTCTCCCAAGTCTCCCAAGAGGTATC 2759
QY 2649 CCGGCTGACAGAGATCAGAGATGACCTGAGCGCTTCCGCAATTCGTTGAGCGCA 2708
DB 2760 TCTG---GAGCAGAGATCAGAGATGATCTCGGGGTGTCCGAGCTTCTGAGGGGA 2816
QY 2709 GGGCCCGCAACCGGCAACAAATGATGACAGTGTGCTGGAACCTGAGAAATTT 2768
DB 2817 TGCCAGACACCGGCGCAACCAATGACAGTGTGTGTGTACCGAACAACAGCTGT 2876
QY 2769 CATGTGACTTCAGCATGCAAGTGAAGATCAATGTTGTGAAACCTGAGAAATTT 2828
DB 2877 CAGGGCAGACCAAGTGAAGCTGTGAGAAACAGCTTTCAGGATATCTGTAAAGAAAT 2936
QY 2829 CAAAAACAGAACGGGTGAGAAAGTGTGTGTGTGTTCACAAATTTCTGCTGTTCTT 2888
DB 2937 CAAAAACAGTCAATGTGTGAGAAAGCTGTGTGTGTGTGTTCACAAATTTCTGTTCTT 2996
QY 2889 CTACAAATGACACAGAGAAATCATCTTGCAGCTGCTGCTGCTGCTGCTGCTGCT 2948
DB 2997 CTACAAATCTCATGAGATGATCACTCCAGTGGCAGCTTCCGCTGTGCTGCTGCTGCT 3056
QY 2949 CACATCTCCCTCTGAGTCCGAGAACATCCAGAAAGCTACGTGTTCAGCTGCTTCAA 3008
DB 3057 GAGCATCTCCAGAGAGCGGATGAGCATACAAAGCTATGTTTTCAGCTTCAAGTTCAA 3116
QY 3009 GTCCCACTCTACTACTTCAAGCGGAAAGCGAGTACAGCTTGAAGAGTGAAGT 3068
DB 3117 ATCCCAAGTCTACTTCTTCCGGCTGAGAGCAAGTACAAATTTGAAAGGTGATGAGGT 3176
QY 3069 GATCGCAGTGCACACAGCTGTGCTGCGACCCC 3103
DB 3177 GATCCAGGGGGCCAGACCTCAGCGGGAGGGGCC 3211

RESULT 12
AB018336
LOCUS AB018336 3997 bp mRNA linear PRI 10-JAN-2004
DEFINITION Homo sapiens mRNA for KIAA0793 protein, partial cds.
ACCESSION AB018336
VERSION AB018336.1 GI:3882306
KEYWORDS
SOURCE Homo sapiens (human)

ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE	1
AUTHORS	Nagase,T., Ishikawa,K., Suyama,M., Kikuno,R., Miyajima,N., Tanaka,A., Kotani,H., Nomura,N. and Ohara,O.
TITLE	Prediction of the coding sequences of unidentified human genes. XI. The complete sequences of 100 new cDNA clones from brain which code for large proteins in vitro
JOURNAL	DNA Res. 5 (5), 277-286 (1998)
PUBMED	9872452
REFERENCE	2 (bases 1 to 3997)
AUTHORS	Ohara,O., Suyama,M., Nagase,T., Ishikawa,K. and Kikuno,R.
TITLE	Direct Submission
JOURNAL	Submitted (08-OCT-1998) Osamu Ohara, Kazusa DNA Research Institute, Laboratory of DNA Technology, Yama 1532-3, Kisarazu, Chiba 292-0012, Japan (E-mail:cdnainfo@kazusa.or.jp, Tel:+61-438-52-3913, Fax:+61-438-52-3914)
FEATURES	Location/Qualifiers
SOURCE	1..3997

```
gene      1. .3997
CDS      /gene="KIAA0793"
          <115. .3282
```

```

/Note="Start codon is not identified"
/codon_start=1
/product="BAA0793 protein"
/protein_id="BAA04513.2"
/db_xref="GI:40788367"
/tranaction="RMEBIGCTATVLOTGAGAGTAPGVSTETGEPOTYLLPRMOEKHKE
LHAEVLLDNTMTEPDIPECDKQVLLTQVWRKLNVEDVDGSEFQNTQSYLLMEK
MKPIIRQRRPKNVVLAVKFFPPDPQQLQETRYTLPALQKDLLEELTADDTT
AALLTSLDSEIGDTELDRLHKVNETLPEQCHCELEKTLPEHQVIGTQPVKSD
VALLTAKLEKLEPFPFMAIDRGRTQLQAVSHMGLVFGYTKTLNTNMAKRLSE
KRKKPILLKHPENVGQVDTLELLESDRECQNFMTI CSEYNTFLTDQCPKPAKV
PFSSEFRRFSGRQKQVUDLPKOSKMTLPERKSRKTLFRLDLPKOSLSEF
EGATPASPASNAFPLVSGTLVSGSLPEPQSDKSLSDPOVSVKSPABEESGA
AGGDPAPASAPLAPLAPSPGTLVSGSPQSPKSSLSLSPAOVLLPBAEGSSAP
LLVSLVSDAGAGMDCEPEPKHVPADENAFYVKELTATERTYLLQDLVTVPRSAV
LVEEDAMPAATVTLFSNIDTYLHNGFLEVEQVLAWEESKATYTGSHQRIGDILL
LNNRQOKETSIFYORHIDVLELAKATRCCKLNAVVEFELQVCLPNTLPLKTE
QORLLAHLRLRLRLCHGVSGHNDVADCDALATLVEYTLQHLITSLGVNGTSP
ORDLVGENTLAPRCEPIREGSCAHKTLKQDLQOMFPLSDMLVTSKAGVSGHPR
KGLLLPGLQMGVYESDNNMSPHCFETLYAAOKTIVVAASRLEKEKMDLNSALQAK
SGCQTALALGVRYCTPRPSPRNVSLPESKEDSDKAVRSSLGEGGRANTATHVCC
YRNTSVSRHSHSAVENQSLGYLLRKHSHQKQKLVVFNFLCFKTHQDDYPLAASH
SLPLGVSQYIPRAGDINHVDYQKPLFSHYHPPAEKCYFERMEVVDIQASSSAG
LAPSLVDDGQPSGSGLEKVMRGEKE"

```

ORIGIN

Query Match	33.1%;	Score 1037.4;	DB 5;	Length 3997;
Best Local Similarity	60.2%;	Pred. No. 1.1e-280;		
Matches 1864;	Conservative	0;	Mismatches 186;	Indels 45;
				Gaps 7;

QY 21 GCCGACCCCGAGATCAAGACTGGGGGGCCCCGGAAAATTGGGGATCAATCTTGAACG 80
Db 150 GCGAGCTGCGGGGATCGCTTGGGTGCCAGACCCTGTGGAGTTACACCTCTTGACC 209
QY 81 TGGACGAAAGCCGCCCCCAACCTTCAGAAAACCTGTCATCAAAATCCAGATGCT 140
Db 210 TGGGCAATCTCTTGGCCAGATGAGAGAGACCTGCACCTCAAGATAAAGCTGCT 266
QY 141 GATGACACCCGAGAGGCAATTGAATTCCACAAAGAGCTCTGGGAGGTGCTGTGA 200

Db	270	GGACAACACCATGGAATATTTGACATTGAGCCTAAATGCGATGGCCAGATTAATCTAC	329
Qy	201	TGCAGTTTGCACCAACCTCAACCTCTGTGGAAGGTGACTATTTTGGCCTCGAGTTTCTCTGA	260
Db	330	ACAAGTGTGGAAAGCGTTTAAACCTGGATGAAATGTGACTACTCTGGGAATGAGGTTTCAAAA	389
Qy	261	TCACAAAAGATCAACGGTGTGGTGGATCTCTCTTAAACCCATTTGTGAACAGATTAGAAG	320
Db	390	TACTCAGTCTACTGGAATTTGGCTTTGAACCTATGAACCCATCTATTTAGGCAAAATACGAAG	449
Qy	321	GCCAAAGCAGCTGTGTTTAAAGTTTGTGTGAATTTCTTTCCGCTGACCAACAACAACCT	380
Db	450	GCCAAAGATGTGTGCTTCCCTAGCCTGTAAATTTTTCACCTGATCTCTGTACACT	509
Qy	381	CCAAAGAAATCACAAGGTACCTGTTCCGGCTGACAGGTGAAGACAGGACTTGGCTCAAG	440
Db	510	ACAAAGAAATATACAAAGATATCTGTTTGGCTTGCACACTTAAGAGAGACTGCTGGAAAGA	569
Qy	441	CAGGTTGACGTGTATATGACACCAAGCGAGCTCTCTTGATTTTACACATTTGTGCAATCTGA	500
Db	570	GCGTTTGAACGTGTCTGTGACACCAAGGGGCCCTTCTCAGCTCCCATCTCTGCAGTCCGA	629
Qy	501	GATTGGGGAATTTTATATGAGCCTTGGACAGAGACACTTACGCAAAAAATTAATACATACC	560
Db	630	AATAGGAATATACATGTAAAGCTGTGACCCGAGACACTCAAAATGAACGATATTTGGCC	689
Qy	561	TCAGCAAGACGCATGAGGACAAATTCGTGGAATTTTCAACATTAACCATTTGACAAAC	620
Db	690	TGGCCAGAGAGACTGCCCTTGAAGAAATCTAGATTTCCATAGAAAGACGTTGGCCAGAC	749
Qy	621	ACCAGCAGATCAAGATTTCCAGCTTCTAGAGATTTGCCCTGGCTAGAGATGTATGAAAT	680
Db	750	ACCTGCTGAGTCGATTTTCCAGGTGCTCGAATTTGCTGAAAGTTTGGAAATGTACGGCAT	809
Qy	681	CCGGTTGACCCGGGCCAAAGACAGGGAAGGACGAAGTCAATCTGGCCGTTGGCAACAC	740
Db	810	CAGATTTTCACTGGTTTCTGACAGAGAAAGAACAAAGTTTAACTGGCAGTTTCCACAT	869
Qy	741	GGGAATTTCTAGTGTTTCAAGGTTTCACTAAGATCAATGCTTCAACTGGGCAAGGTGCG	800
Db	870	GGGTGTACTCGTGTTCAGGGCACAACAATAATCAACATTTCAACTGTGTCGAAGTCCG	929
Qy	801	GAACTGAGCTTTCAAAGGAAGGCGCTTTCTCATCAAGTCCGGCCAGATGTCATATGTC	860
Db	930	TAAACTAAGCTTCAAGGAAAAAATTTCTTATCAAACTTATCAGAGGTTCAATGAGCC	989
Qy	861	GTACCAAGATTCCTTGGATTCCTGATGCGCAGTCGGGATTTCTGCAAGTCTCTTGGAA	920
Db	990	TTTACCAAGACACATTAAGATTTTGTGGGTATGATAGATGAATGTAAAGAACTTCTGGA	1049
Qy	921	AATCTGTGTGAACATCATGCTCTTCTTTAGACTTTTGAAGAGCCCAAAACCAAGCCCA	980
Db	1050	GATTTGTGTGAGTATCACACTTTTAAAGCTTTTGGACCAACCTTAAGCCAAAAGCAAA	1109
Qy	981	GCCGTCCTCTTTAGCCGGGGGTATCATATTTCCGTTCAAGTGTGCGAATCTGAAGCAGAT	1040
Db	1110	AGCGGTCTTCTTCAGCCGGGCGCTCTCTTCAAGATCAAGTGAAGAACTCAGAAACAAT	1169
Qy	1041	TCTCGACTATTTTAAAGAGGAGACATTAAGAGGTGCACTTTGAAGGAAGACACAGCA	1100
Db	1170	AGTAGATTAATTTCAAAAGCAGGTGAATGAAGAAATTCATATGAAGAAAGGACACAGCA	1229
Qy	1101	GATTCATTTCTATCCGGAGCCTTGCTTACAGCCTTACAGAACTGAATGAATTCGGAAGTCTGGA	1160
Db	1230	GACCCACACGTCC-----GTTTCGAGCTTGACTGAGACTTACAAA	1271
Qy	1161	GCAGTCTCAGCAGACACAGCCTTATACATTTGGAAGAGGTGCGCAATTTCCAGGGGCGCA	1220
Db	1272	ACAGAGACTCTCATTTCCCGAGGGATTTAGGAGCTCTGCTCCCATCTTTCAGGAAGATGC	1331
Qy	1221	GAGCTCCCGCGAGAAAAGAACCGAAGTTTCCGCGGGGAGCCGGGGTCTGCACCCGAG	1280
Db	1332	CTTTTACTCTGCTCTCTCTTCACTCTGTGTCCTCTTGGCCCTGCGCAGGTTTAAAGACAG	1391

Scheetz, T.E., Brownstein, M.J., Uedin, T.B., Toshiyuki, S., Carinci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahy, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Rodchenko, Y., Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smal, D.E., Scherzer, A., Schein, J.E., Jones, S.J., and Marra, M.A. Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

JOURNAL
PUBMED
12477932

2 (bases 1 to 3719)

Strausberg, R.
Direct Submission
Submitted (05-JUN-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: amg@bcm.tmc.edu
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Louisse, H., Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati, A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
Series: IRK Plate: 5 Row: m Column: 20
This clone was rejected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 21704021.
Location/Qualifiers
1. 3719
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N-3"
/db_xref="taxon:10090"
/clone="MGC:6304 IMAGE:2655209"
/tissue_type="Mammary tumor. MMTV-LTR/INT3 model. 5 month old mouse. Taken by biopsy."
/clone_11b="NCI CGAP Mam2"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
1. 3719
/gene="Farp2"
/note="synonym: Flr"
/db_xref="GeneID:227377"
/db_xref="MGI:2385126"
99. 3296
/gene="Farp2"
/codon_start=1
/product="ERM, Rhogef and plectstrin domain protein 2"
/protein_id="AAH09153.1"
/db_xref="GI:14318719"
/db_xref="GeneID:227377"
/db_xref="MGI:2385126"
/translation="MGEIGTYRALPTSGTRIGGOTAIQVSTLEPQSLSPMOEKHM RVRKLDSTVELFDIEPKDGVLLITVYKHLNIBCDYFGIEKPNVQSVYIMLEPM KRIIOVRKNAVIRLAKFPDPDGOIEBYTRVLPDLCRDLERLRLCTNNTA ALIISHLQSEIGDDETDREHLKANEYLIPNOEKSLKILDFHQHHTGQTAEEDFO VLEIARKLEMYGIRFHMASDRGTKINILAVSHGVLVFQGTIKINFPMSKVRKLSFK

FEATURES
source

gene

CDS

ORIGIN
Query Match 32.3%; Score 1013.2; DB: 6; Length 3719;
Best Local Similarity 59.3%; Pred. No. 7.4e-274;
Matches 1846; Conservative 0; Mismatches 1238; Indels 30; Gaps 6;
22 CCGACCCGAGATCAAGACTGGGGGCCCGGAAATTCGGGATCAGTACCTTGAACGT 81
132 CCAACTTCAAGGACACGCTGGGTGGCCAGACCGCATTTGAGTACCTCGAGACCA 191
82 GGACGAAGCGCGCCCAACACTTCAGAAATCTGTTCATCAATAATCCAGTCTG 141
192 GAGCAGACCTGTCAACCCAGATGAGAGACATGCGCATCAGATGAAGCTGCTG 251
142 GATGACACCCAGAGGACATTTGAATTCACAAAGAGCTCCGGGAAGTGTGCTGAT 201
252 GACAGCAGTGAAGTATTGATGACATGAGCCGAAGTGAACGGGACAGCTTACTGACT 311
202 GCAATTTGCAACCACTCAACTCGTGAAGGTGACTATTTTGGCTCGAGTTCTGAT 261
312 CAAGTGTGAAGCATTTGATATCTCATGATGTGACTATTGTGTGAGTTCAAGAT 371
262 CACAAAAGATCACGGTGTGCTGATCTCTAAACCATTGTGAACAGATTGAAGG 321
372 GTCCAGTCTCACTGATTTGGCTTGAACCTATGAACCATATATAGCAATGAGAAG 431
322 CCAAGACAGTGTGTTGAATTTGTTGTTGTAATCTTTCGCGTCGACACACACACTC 381
432 CCAAAAATGGGCTCTCGCGACGATTAATTTTCCGCGCTGACCTGCTGACTG 491
382 CAGAAGAACTCAAGAGTACCTGTCGCGCTGACAGTGAAGCAGACCTTGGCTCAAGGC 441
492 CAGAAGAGTACACAGAGTACCTGTTGCTTGAACCTCAAGAGAGACCTCTGAGGAA 551
442 AGTTGACGTGTATGACACAGCGACGCTCTTGAATTTACACATTTGTCATGTGAG 501
552 CGCTGACCTGACCGGCAACACTGACAGCCCTTCTCATATCCACCTTCTGACGTGGA 611
502 ATTGGGGAATTTGATGAAGCCTTGGACAGAGCACTTGAAGAAATTAATATACCT 561
612 ATCGAGATTTATGATGAACCTGATGAGAACCTTCAAGCCAAATGAATATCTGCCC 671
562 CAGCAAGACGACTAGAGAGCAAAATCTGGAATTTCAACATTAACCAATTGACAAACA 621
672 AACGAGGAATCTCTGAAGAAAGTACTAGACTTCAATGAGAGCACACGGGCCAGAT 731
622 CCAGCAATCAGATTTCCAGTCTTGAAGATTTGCGCGCTGAGAGATGATGATGATC 681
732 CCGTGAAGTCAAGATTTCCAGTCTTGAAGATTTGAGATTTGAGATTTGAGATTT 791
682 CCGTTCACCGGCGCAAGAGAGGAGGACGAGATCAATCTGCGCTGCGCAAGACG 741
792 AGTTTCAATGAGCTTCTGACAGAGAGAGGACGAGATTTAATCTGACGTTTCTACATG 851
742 GGAATTTAGTGTTCAGGCTTTCATGAATGATGATGCTTCACTGCGGCGAGGTGCG 801
852 GGTGCTGCTGTTCAGAGGTACACGAAATCATATCTTCAATGCTGCAAAAGTCGT 911
802 AAGTGAAGCTTCAAGAGAGAGGCTTTCTCATCAAGCTTCGCGGCGAGATGCCAATGATGCG 861

Db 912 AAACTAAGCTTCAAGAGAAAAGTTTCTTATCAAACTCCACCCTGAGGCTCATGAGCC 971
Qy 862 TACCAAGATACCTTGAATTCCTGATGGCCAGTCGGATTTCTGCAAGCTCTTGGAAA 921
Db 972 TACCAAGACCGTTAGAGTTCTTGTGGTAGAGAGATGAATGTAAAGCTTCTGAAAG 1031
Qy 922 ATCTGTGTGAACATCATGCTCTTGTAGACTTTTGAAGAGCCCAACCAAGCCCAAG 981
Db 1032 ATATGTGTGAAGACATACCTTTTAACTCTTGAAGCTTAAAGCTTAAAGCCAAAG 1091
Qy 982 CCGGCTCTTTAGCGGGGGTCAATTCGGTTACAGTGTGGAATCAAGAGAGTT 1041
Db 1092 GCTGTCTTTCAGCGAGGCTCTCTCTTCAATACAGTGAAGAACTCAAGAAACAATA 1151
Qy 1042 CTCGACTATGTTAAAGAGAGACATTAAGAGTGCATTTAAAGAAAGACAGCAAG 1101
Db 1152 GTAGATTATGTCAAAACGGTGTGAATGAAGAAATTCATACAGAAAGACAGGTAAAG 1211
Qy 1102 ATTC---ATTCTATCCGAGCTTGTCTTCAAGCTTACAGAACTGAATTGGAAGTGTG 1158
Db 1212 ACTCGAATCTTTATATGCTGTGACTGTAGATCTGCTTAAACAGAGCTCTTCAAC 1271
Qy 1159 GAGCAGTCTCAAGCAGACACAGCTTCAATTTGAAGAGTCCGAAATCTCAAGGGGG 1218
Db 1272 GATGGCTTGAAACTTCTGCTCTCTGTCTTCAAGAAATGTCTCTTTATTC----- 1324
Qy 1219 CAGAGTCCCGGAGAGAAAGAAACGGAAGTTTCCGCGGGAGACCGGGGTGCACCG 1278
Db 1325 ---ACCCCTTAGTCTTCTTGTCTCTCTCTGTCTGCTTGAAGAAAGACAGCAGAG 1381
Qy 1279 AGCCCTGCGCGAGAGAGAGCCCGGGGTAAACAAGAGCGGACGGAAGCGCTCGAGG 1338
Db 1382 CTCCTTGTGATCCGAGGCTCCGTCATCAAGAGACAGCAGCAGAGAGAGCAGCGG 1441
Qy 1339 CCCACGAGAGAAAGAGAGATCGTTAAGATAGAACCCAGACAGATAACTCAGCCC 1398
Db 1442 ACCA---TCAATCATCTGACCGGCCACAGACAGTCGGCCATCTCCCTGAGCCCCGG 1499
Qy 1399 CCGCAGCCAAAGCAGAGCTCCCTGACTGGCAGTCTCACTTTCCAGAGCTGTGTGAAC 1458
Db 1500 CTCGCGCTGTGTCCAGAGCTTTTCTATGATAGTCTCAAGCTTCTCTCCACAGCCCTGAAG 1559
Qy 1459 TCGCAGGGGGAGTGGCCCTGCGCAACGTGACCTTGTCTCCAACTGAGCCCGCAGACC 1518
Db 1560 AGCCACTGAGCTGTGCTGCTGAGCTTCAAGGCGCG-----CTGAGTACAGCTGAG 1610
Qy 1519 AAGCAGGCTCTTCCCTTGAATCAGCCGCTGTGAATGACCAAGGCTGCCCCGGAAGCAG 1578
Db 1611 CAGGGTGCATCCCTGTGTCTGAGCCCTGTGTCTCAAGTGGCGCTGTACAGCAGAGATGAGC 1670
Qy 1579 GATGAGATGAGGGCCGAGAGAAAGATTCCTCACTGATTAAGCGTACTTCATAGCTTAAG 1638
Db 1671 AACCAAGAGAGAGCAAAACACAGACATGCTCAGAAAGATGAGGCTATTTTCAATAGCAAG 1730
Qy 1639 GAAGGTCTAACACCGAGCGAATATCTGAAGAGATCTGGAAGTATCACTTCTGTGTTT 1698
Db 1731 GGAATTTCTGCTACAAACGAACTTATCTGAAGATTTAGAAATATCACTGTGTGTTT 1790
Qy 1699 CAGAGCAGATGAGCAAGAGAGAGCAGCATGCTCGGAAGCACTGAAGATCTCATATTTCCG 1758
Db 1791 CGGAGGTGTCTATCAAGAGAGAGGACATGCTCGAGCCCTGATGAGCCCTGTTTCTCC 1850
Qy 1759 AATTTGAACCTTTGCACAAAATTTCACTAATTTTCTCAAGGAAATGAGACAGACTT 1818
Db 1851 AACATTGATCTCAAGTCTTACAGATTCACAGAGCTTCTTCAAGAGTGTGAACAGAGCTG 1910
Qy 1819 GCCCTGTGGAAGAGCGCTCAATGCGCA---AATCAGAGATTAACAAAGATCGCGAT 1875
Db 1911 GCACTTGGAGAGGCGCTTCAAGTGCCTTAAAGATGATACCAAGCGAATCGGAGAC 1970
Qy 1876 GTCATGCTGAAGAACTTCAAGGAGTGAAGCAGCTGCGGCTCACTGTGGAAGCAGC 1935

Db 1971 ATCTCTCAGGAACATGCGTCAAGTTAAAGAAATTTACTAGCTACTTCCAAAGACAGAT 2030
Qy 1936 GAGGCTTGAAGGCGCTGAGAGATGGAATCAAGAGCTCCCGGCGGTGAGAACTTCTGC 1995
Db 2031 GAGGTCTTAAACGAATCTGGAAGAGCCACAAACACTGTAAAGATTTGAGGAGTCTAC 2090
Qy 1996 AGAGCTTTGAGCTGAGAAAGGTGTGTTACTACCGCTCAACACTTCTCTGCGGCA 2055
Db 2091 AAAGATTTGAGTCCAAAGAGTCTGTCTCTCTCTCAACACTTCTCTGAGAGCC 2150
Qy 2056 CTGCACCGCTCATGACATCAAGCAGAGTCTGGAAGGCTGTGCAACACACCGCGG 2115
Db 2151 GTCCAAAGGCTAGTCCACTACCGTGTGTGAGCGGCTGTGTCTCACTCTCTCT 2210
Qy 2116 AGCCACCGGACTTCAAGGACCTGCGAGCGCTTTGAGAGATCAAGGATGTGGCA 2175
Db 2211 GGGCACCGGACTATGCCCATGTCATGAGGCACTAAGGCCATCAAGAGTGCACACC 2270
Qy 2176 CAGCTCACGGTACATGATCAAGATGAGAAATTTCCAGAGCTGCACGAACTCAAGAA 2235
Db 2271 GAGCTCACGAAGAGCTTACCGGCTGGAAGAACTTCAAGAAATTTGAGAGTACAGCA 2330
Qy 2236 GATTGATTTGCAATGACAAATCTTGTGTTCCGGGAAAGGAGTTCAATCCGTCTGGCAGC 2295
Db 2331 GACTGTGCTGTGTAAAGAACTCATTTGCTCTGAGGAGGAGTTTATCGTGAAGGCTGC 2390
Qy 2296 CTCAGCAAGCTCTCGGGGAGAGGGGCTCAGAGCGCATGTTCTCTCTTCAACAGAGTC 2355
Db 2391 CTGCACAACTCAACAAAGAGGCTGCAACAGAGATGTTTTTCTGTCTCAGATAG 2450
Qy 2356 CTGCTATACAGAGCGGGGGCTGACGCGCTCCAACTGATTTAAAGTCCACGGGAGCTC 2415
Db 2451 TTACTGTATACAAAGAAAGTGTCAAGAGAGCCAGTCAATTCGGATTCGTGCTTCTT 2510
Qy 2416 CCGCTTATGCAATGCAATGAGAGAGCGAAGCAGATGAGGGGTGCCCATCTGCTG 2475
Db 2511 CCACTCGTGTGATGTGTAGAAAGATGAAATGAAATGATGCTTCTTCAATGCTTC 2570
Qy 2476 ACCCTCGGGGCGAGGGGAGTCCATCATGTGCGCGCAATTCGTGCTCCAGATGAG 2535
Db 2571 ACCATTTAGAGCTCAGAAACAAATGTGTAGAGCAGCAGCTGGCTAGAAAGAA 2630
Qy 2536 AAGTGGTTGAGGACATCCAGATGCTTGAACCTTGGCGAGAGAGAGCAGAGCCCGCT 2595
Db 2631 AAGTGAATGAGGACCTGTAATGACAGATTCAGAGCAGCAAGACTATCGGTACTACCC 2690
Qy 2691 CAGGTCTGTGG---GAGGCGGGGTATACGTGACCTTATGATCTTCTGATGAATC 2747
Db 2656 GACCAAGAGTCAAGAGATGACTGAGCGCTGCGCAATCGTGTGAAGCGCCAGCCCG 2715
Qy 2748 TCTCTGGAAGAAATCAGAAATGTCAGAGAAACCGGGGCTCTCTGAGGGGAAACAGCCAG 2807
Qy 2716 CACCGGCGCAACCAATGCTGACGATGCTGCAACCGCAACCGAGCTCTCATAGTGTG 2775
Db 2808 CACCGGCGCAATTAACCAATGATGTGTCTGTACCTGTATTAACAGTGTCTCAAGCA 2867
Qy 2867 GACTTCAGATCGCAGTGAAGATCAATGTGTGGAACCTGTGTGAAGAAATTCAGAAAC 2835
Db 2868 GACCAAGTGAAGCTGTGTAAGACAGCTTTCAGGATATCTGTGTGAAGAAATTCAGAAAC 2927
Qy 2836 AGCAACGGGTGGCAAGAGTGTGGGTGTTCACAAACTTCTGCTGTCTTCTTCAAA 2895
Db 2928 AGTAATGGCTGCAAGAGCTGTGGGTGTCTTTCACAACTTCTGCTGTCTTCTTCAAA 2987
Qy 2896 TCACACAGAGCAATCATCCCTTGCAGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCT 2955
Db 2988 ACACATCAGAGATGATCAACCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3047
Qy 2956 CCCTCTGAGTCCGAGAACTCAGAAAGACTAGTGTCAAGCTGACTTCAAGTCCAC 3015
Db 3048 CCAGGAGGCTGACAGATCAACAAAGACTATGTCTTCAAGCTCAATTCAAATCTCAC 3107


```

QY      3016 GTCTACTACTTCGAGGCGGAAACGAGTACACGTTTGGAAAGTGATGAGATGATCCGC 3075
DB      3108 GTCTACTTCTTCGCGGCTGAGAGCAGTACACTTTGAAAGGATGATGACGTCATCAAA 3167
QY      3076 AGTCCACCAAGCTCTGCTCGCGAGCCACCGTGTAGCCACAAAGATCTCTT 3129
DB      3169 AGGGCCAGCAGCTCAGCAGGAGAGACCCCAAGTTTCACTCAGGACTGCTCACAT 3221

RESULT 14
LOCUS   BC027077
DEFINITION Mus musculus FERM, Rhogef (Arhef) and pleckstrin domain protein 1
(chondrocyte-derived), mRNA (cDNA clone IMAGE:5010682), partial
cds.
ACCESSION BC027077 2431 bp mRNA linear ROD 21-OCT-2003
VERSION BC027077.1 GI:20071584
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridea; Muridae; Murinae; Mus.
1 (bases 1 to 2431)
Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschuler,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Datchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
Carinci,P., Prange,C., Raha,S.S., Lochlano,N.A., Peters,G.J.,
Abramson,R.D., Millay,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richard,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J., Helton,E., Kettaman,M., Madan,A., Rodrigues,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakeley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalins,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 2431)
Strausberg,R.
Direct Submission
Submitted (04-APR-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgabbs-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
cDNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: amg@bcm.tmc.edu
Gunnar, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Louise, H.,
Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Navati,
A.N., Gibbs, R.A.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov
Series: IRAC Plate: 45 Row: C Column: 12
This clone was selected for full length sequencing because it

```

passed the following selection criteria: Hexamer frequency ORF analysis.

FEATURES

source

Location/Qualifiers

1..2431

/organism="Mus musculus"

/mol_type="mRNA"

/strain="FVB/N-3"

/db_xref="taxon:10090"

/clone="IMAGE:5010682"

/tissue_type="Mammary tumor, MMTV-LTR/INT3 model, 5 month

old mouse. Taken by biopsy."

/clone_lib="NCI CGAP_Mam2"

/lab_host="DH10B"

/note="vector: pCMV-Sport6"

<1..2431

/gene="Farp1"

/note="synonym: Cdepr"

/db_xref="GeneID:223254"

/db_xref="MGI:2446173"

<1..1065

/gene="Farp1"

/codon_start=1

/product="Farp1 protein"

/protein_id="AAH27077.1"

/db_xref="GI:20071585"

/db_xref="GeneID:223254"

/db_xref="MGI:2446173"

/translation="MKHLERLCKHPNHPADFRCPALALATETVVAQHGTMKME

NPKHKEIKKLDIGDVTGPERIRIGSLKSGKLGOMFLPNDVLLYTRGL

TRSNPKYHGLPLKMTIESEBEMGCPHCTITRGQROSTIVAASSREBEMKEDI

QWADLAKNSGPELILASSPDDKSPDEATPAQDSBDLSASRTSLERQPHKGN

IMVHVCWRSTSVSVDPSIAYENDLNRKFNNSGWMKLVWFTNFCIFPKSH

QDSHPILASPLILGSLIPSESEINHKDYVFKLHKSHYVYFRAESEYFERMEVIR

SATSSASRAHLISKESHLTY"

250..501

/gene="Farp1"

/note="PH; Region: Pleckstrin homology domain. Domain

commonly found in eukaryotic signalling proteins. The

domain family possesses multiple functions including the

abilities to bind inositol phosphates, and various

proteins. PH domains have been found to possess inserted

domains (such as in PLC gamma, synrophin) and to be

inserted within other domains. Mutations in Brucins

tyrosine kinase (Btk) within its PH domain cause X-linked

agammaglobulinemia (XLA) in patients. Point mutations

cluster into the positively charged end of the molecule

around the predicted binding site for phosphatidylinositol

lipids"

/db_xref="CDD:smart00233"

730..1020

/gene="Farp1"

/note="PH; Region: Pleckstrin homology domain. Domain

commonly found in eukaryotic signalling proteins. The

domain family possesses multiple functions including the

abilities to bind inositol phosphates, and various

proteins. PH domains have been found to possess inserted

domains (such as in PLC gamma, synrophin) and to be

inserted within other domains. Mutations in Brucins

tyrosine kinase (Btk) within its PH domain cause X-linked

agammaglobulinemia (XLA) in patients. Point mutations

cluster into the positively charged end of the molecule

around the predicted binding site for phosphatidylinositol

lipids"

/db_xref="CDD:smart00233"

ORIGIN

Query Match

Best Local Similarity 89.3%; Pred. No. 76-236;

Matches 947; Conservative 0; Mismatches 114; Indels 0; Gaps 0;

QY 2074 TCACAGCAGTCTCGAGCGGCTGTGCAACACCCGCGAGCCAGCCGACTTCAGG 2133

DB 1 TATAAGCAGTCTCGAGCGGCTGTGCAACACCCGCGAGCCAGCCGACTTCAGG 60

QY	2134	GA	CTCCGAGCCGCTTTGGCAGAGATCAACGGAGATGGTGGCAGCACTCCAGCTCACGAT	2133
Db	61	GA	CTCAGAGCTCCGCTGGCGGAGATCAGAGATGGTGGCAGCACTCGATGGTCCATG	120
QY	2194	AT	CAAGATGAGAAATTTCCAGAAAGTCGACGAACCTCAAGAAAGATTGATGGCAATTGAC	2253
Db	121	AT	CAAGATGAGAAATTTCCAGAAAGTCGACGAACCTCAAGAAAGATTGATGGCAATTGAC	180
QY	2254	AAT	CTTGTGCTTCGCGGAAGGGAGTTCATCCGTCTGGGCAAGCTTCAGCAAGCTCTCCGG	2313
Db	181	AAT	CTTGTGACCCCAAGGAAGGAGTTCATCCGCTGGGCAAGCTTCAGCAAGCTCTCCGG	240
QY	2314	AAG	GGCTCCAGCAGGCAATGTTCTTCCGTTCAACGAGTCGCTTACACAGCGCG	2373
Db	241	AAG	GGCTTCAGCAGCGCAATGTTCTTCCGTTCAACGAGTCGCTTACACAGCGCG	300
QY	2374	GCG	CTGACGCGCTCCAAATCAGTTTAAAGTCCAAGGCAAGCTCCGCTTATGCAATGACG	2433
Db	301	GGA	CTGACGCGCAATCAGTTTAAAGTCCAAGGCAAGCTCCGCTTATGCAATGACG	360
QY	2434	ATT	GAGAGAGAGGAAGACGATGGGGGGTCCCACTGCTGACCTTCCGGGCGACGCG	2493
Db	361	AT	CGAGAGAGGAGAGAGAGTGGGTGTGCCCACTGCACTTCGCGGGCGACGCG	420
QY	2494	CAG	TCCATCATGATGAGCGCCGACGTTCTGAGTCCGATGGAAGATGGATTTGAGGACATC	2553
Db	421	CA	ATTCATCATGATGAGCTGCTGCCAGTTCCGGTCAAGATGGAAGATGGAATGGAACATC	480
QY	2554	CAG	ATGACCATTTGACCTGGCGGAGAGACAGCAGCCCCGCGCTGATTCCTGCGCAGC	2613
Db	481	CAG	ATGACCATTTGACCTGGCGGAGAGAGACAGCAGCCCCGCGCTGATTCCTGCGCAGC	540
QY	2614	AG	CCCCCTGACAAACAATGCCCTTATGAAGCAGCCGCGCTGACAGAGATCAGAGAT	2673
Db	541	AG	CCCCCTGACAAACAATGCCCTTATGAAGCAGCCGCGCTGACAGAGATCAGAGAT	600
QY	2674	GAC	CTGACGCGCTCGCGCAATCGGTGAGCGCCAGGCGCCGACCGCGGAAACAATG	2733
Db	601	GAC	CTGACGCGCTCTCGGACATCGTGGAGCGCCAGGCGCTTCACCGTGGCAACATG	660
QY	2734	GTC	ACGTGTCTGGCACCGCAACAACGAGCTTTCATGATGAACTTCAGCATCGCAATG	2793
Db	661	GTC	ACGTGTCTGGCATCGCAGCACCAAGTGTCTCATGTGTGAACTTCAGCATCGCAATG	720
QY	2794	GAG	AATCAGTTGTCTGGAACCTGTGGAATTCAAAACAGCAACGGGTGGCAAG	2853
Db	721	GAG	AATCAGCTCTCTGGGAATTCGTGGAAGATTCAAGAACAGCAATGGCTGGCAGAG	780
QY	2854	CTG	TGGGAGTGTTCACAAACTTTCGCTGTTCTTCTAACAATCAACACAGAAATCAT	2913
Db	781	CTG	TGGGAGTGTTCACCAACTTTCGCTGTTCTTCTAACAATCAACACAGAAATCAT	840
QY	2914	CCC	TTGCGACCTGCTCTGCTCGGCTACTCGTCAACATCCCTCTGATTCGAGAAC	2973
Db	841	CCC	TTGCGACCTGCTCTGCTCGGCTACTCGTCAACATCCCTCTGATTCGAGAAC	900
QY	2974	AT	CCAGAAAGCTTACGTTTCAAGTGTGCACTTCAAGTCCCAAGTCTACTTACAGGCG	3033
Db	901	AT	CCAGCAAGACTTATGTTTCAAGCTGATTTCAAGTCCCAAGTCTACTTACAGGCT	960
QY	3034	GAA	AGCAGTACAGTTTGAAGAGTGAATGATCCGCAAGTCCACAGGCTTGGC	3093
Db	961	GAA	AGCAGTACACTTTGAAGAGTGAATGATCCGCAAGTCCCAAGGCTTGGC	1020
QY	3094	TGC	GACCCACGTTGTGAGCCACAAGAGTCTCTTGTGTA	3134
Db	1021	TCT	GAGGCCACATCTTGATCACAAAGAAATCTCACCTGTA	1061

DEFINITION	Mus musculus FERN, RhoeF (A78ge) and plackerin domain protein 1 (chondrocyte-derived), mRNA (CDNA clone IMAGE:5374219), partial cds.
ACCESSION	BC043327
VERSION	BC043327.1 GI:28175181
KEYWORDS	
SOURCE	
ORGANISM	Mus musculus (house mouse)
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 2296)
AUTHORS	Strauberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,I., Shennan,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetlow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Heish,F., Ditchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stepien,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheer,T.E., Brownstein,M.J., Uddin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Kaha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullaly,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Wallez,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalón,D.K., Muzny,D.M., Sodegren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Kettelman,M., Madan,A., Rodriguez,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schwartz,J., Myers,R.M., Butcherfield,Y.S., Krzywinski,M.I., Skalski,U., Smalls,D.E., Scherach,A., Schein,J.E., Jones,S.U. and Marra,M.A.
TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
PUBMED	12477932
REFERENCE	2 (bases 1 to 2296)
AUTHORS	Strauberg,R.
TITLE	Direct Submission
JOURNAL	Submitted (09-JUN-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT	Contact: MGC help desk Email: gcaps-remail.nih.gov Tissue Procurement: Gilbert Smith, Ph.D. cDNA Library Preparation: Life Technologies, Inc. cDNA library Arrayed by: The I.M.A.G.E. Consortium (LNLB) DNA Sequencing by: Genome Sequence Centre, BC Cancer Agency, Vancouver, BC, Canada info@bcgc.bc.ca Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butcherfield, Susanna Chan, Readman Chin, Chris Fjell, Erin Garland, Ran Guin, Letticia Hisiao, Martin Kirzywinski, Reta Kutische, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mahewson, Katherine McLeavy, Steven Nees, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saadedi, Jacqueline Schein, Duane Smalls, Michael Smith, Lorraine Spence, Jeff Scott, Michael Thorne, Miranada Tsai, Natsija van den Bosch, Jill Vardy, George Yang, Scott Zyderdunyn, Marco Marra.
FEATURES	
SOURCE	Location/Qualifiers 1..2296 /organism="Mus musculus" /mol_type="mRNA" /strain="PVE/N-3" /db_xref="taxon:10090" /clone="IMAGE:5374219" /feature_type="Mammary tumor, MMTV-LTR/INT3 model, 5 month analysis."

old mouse. Taken by biopay."
/clone.lib="NCI CGAP_Mam2"
/lab.host="DHIOB"
/note="Vector: PCMV-SPORT6"
<1..2296
/gene="Farpl"
/note="synonym: Cdep"
/db_xref="GeneID:223254"
/db_xref="MGI:2446173"
<1..926

CDS

/gene="Farpl"
/codon_start=3
/product="Farpl protein"
/protein_id="AAH43327.1"
/db_xref="GI:28175182"
/db_xref="GeneID:223254"
/db_xref="MGI:2446173"
/translation="KLHELKDLIGIDNLVTPGRERIRLSLSKLGKLDQRMPLF
NDVLVTSRGLTASNQFKVHQLPLVGMTESESEBEGVPHCLTRGQROSIIVASS
REMERKMDIOMALDLAEKSNPPELLASPPNDKSRFTAADQESDLSART
SLERQAPHRGNTMVHVCWHRSTSVSNVDSIVENQSLRKFKNSGDKLVVF
TPFCLPFYKSHODSHPLASLPILGYSLTIPSESENIHKDYVKLHFKSHVYFRASE
VTFERMEVIRSRATSSASRAHLHSKESHLV"
111..362
/gene="Farpl"
/note="PH; Region: Pleckstrin homology domain. Domain
commonly found in eukaryotic signalling proteins. The
domain family possesses multiple functions including the
abilities to bind inositol phosphates, and various
proteins. PH domains have been found to possess inserted
domains (such as in PLC gamma, syntrophins) and to be
inserted within other domains. Mutations in Brutons
tyrosine kinase (Btk) within its PH domain cause X-linked
agammaglobulinemia (XLA) in patients. Point mutations
cluster into the positively charged end of the molecule
around the predicted binding site for phosphatidylinositol
lipids"
/db_xref="CD:smart00233"
591..881
/gene="Farpl"
/note="PH; Region: Pleckstrin homology domain. Domain
commonly found in eukaryotic signalling proteins. The
domain family possesses multiple functions including the
abilities to bind inositol phosphates, and various
proteins. PH domains have been found to possess inserted
domains (such as in PLC gamma, syntrophins) and to be
inserted within other domains. Mutations in Brutons
tyrosine kinase (Btk) within its PH domain cause X-linked
agammaglobulinemia (XLA) in patients. Point mutations
cluster into the positively charged end of the molecule
around the predicted binding site for phosphatidylinositol
lipids"
/db_xref="CD:smart00233"

misc_feature

misc_feature

/gene="Farpl"
/note="PH; Region: Pleckstrin homology domain. Domain
commonly found in eukaryotic signalling proteins. The
domain family possesses multiple functions including the
abilities to bind inositol phosphates, and various
proteins. PH domains have been found to possess inserted
domains (such as in PLC gamma, syntrophins) and to be
inserted within other domains. Mutations in Brutons
tyrosine kinase (Btk) within its PH domain cause X-linked
agammaglobulinemia (XLA) in patients. Point mutations
cluster into the positively charged end of the molecule
around the predicted binding site for phosphatidylinositol
lipids"
/db_xref="CD:smart00233"

ORIGIN

Query Match 24.5%; Score 766.8; DB 6; Length 2296;
Best Local Similarity 89.5%; Pred. No. 2.6e-204;
Matches 825; Conservative 0; Mismatches 97; Indels 0; Gaps 0;

QY 2213 AGAAGCTGACGAACTCAAGAAAGATTGATGCAATCTTGTGTTCCGGGAA 2272
DB 1 AGAAGCTGATGAGCTCAAGAAAGATCTGATCGGCATTGACATCTTGTGACCCGAGAA 60
QY 2273 GGAAGTTCACTCGCTGGGGCAGCTCAGCAAGCTTCGGGGAAAGGGGCTCCAGCAGCGCA 2332
DB 61 GGAAGTTCACTCGCTGGGGCAGCTCAGCAAGCTTCGGGGAAAGGGGCTTCAGCAGCGCA 120
QY 2333 TGTTCCTCTGTTCAACGACGTCCTCTATACAGACCGGGGGCTGACGGCTCCCAATC 2392
DB 121 TGTTCCTCTGTTCAACGACGTCCTCTATACAGACCGGGGGCTGACGGCATCTAATC 180
QY 2393 AGTTAAAGTCCAGGGGACGCTCCGCTCTATGCGATGACGATGAGAGAGGAAAGACG 2452
|||||

DB 181 AGTTAAAGTCCAGGACGACGCTCCCACTCTATGCGATGACGATGAGAGAGTGAAGAG 240
QY 2453 AGTGGGGGGTGGCCCACTGCTGACCTCCGGGGGCGCAGCGGAGTCCATCATCTGGGCGG 2512
DB 241 AGTGGGGGTGTGCCCACTGCTGACCTCCGGGGGCGCAGCGGAGTCCATCATCTGGGCTG 300
QY 2513 CCAGTTCTCGGTCGAGATGGAAGAGTGGGTTGAGAGACATCCAGATGAGCATTTGACTTG 2572
DB 301 CCAGTTCCCGGTCAAGATGGAAGAAATGAGATGAGAGACATCCAGATGAGCATTTGACTTG 360
QY 2573 CCGAAGAAAGACGACGACCCCGCCCGCTGAGTTCCTGCGCAGGACCCCGCTGCAACAAGT 2632
DB 361 CAGAGAAAAGCAATGACCCACCCCGAGCTCTTGCCAGCAGCCCCCTGACAAACAGT 420
QY 2633 CCCCTGATGAAGCCACCGCGGCTGACAGAGTCAAGAGATGACCTGAGCGGCTGCGCA 2692
DB 421 CCCCGATGAAACCCACAGCGGCTGACAGAGTCAAGAGATGACCTGAGTGCCTTCGGA 480
QY 2693 CATGCTGAGAGCGCAGGCGCCCGCAGCGCGGCAACAATGATGACGCTGCGCACCC 2752
DB 481 CATGCTGAGAGCGCAGGCGCCCGCAGCGGCAACAATGATGACGCTGCGCACCTGCGCATC 540
QY 2753 GCAACACAGGCTCTCAATGATGACCTCAAGATGCAATGCAAGATCAAGTGTCTGGA 2812
DB 541 GCAGACACAGTGTCTCAATGATGACCTCAAGATGCAATGCAAGTGTCTGGA 600
QY 2813 ACCCTGAGGAATTTCAAAAAACAGCAACGAGTGGCAGAGCTGTGGTGTTCACAA 2872
DB 601 ACTTGTGAGAAAGTTCAAGAACAGCAATGCTGGCAGAGTGTGGTGTTCACAA 660
QY 2873 ACTTGTGCTGTCTTCTTCAAAATCAACACAGCAATGCTGGCAGAGTGTGGTGTTCACAA 2932
DB 661 ACTTGTGCTGTCTTCTTCAAAATCAACACAGCAATGCTGGCAGAGTGTGGTGTTCACAA 720
QY 2933 TGCCTGGCTACCTCGCTCAACCAATCCCTGAGTCCGAGAACATCCAGAAAGCTACGTGT 2992
DB 721 TGCCTGGCTACCTCGCTCAACCAATCCCTGAGTCCGAGAACATCCAGAAAGCTACGTGT 780
QY 2993 TCAAGCTGACTTCAAGTCCCAAGTCTACTACTTCAAGGCGGAAAGCGAGTACAGCTTGC 3052
DB 781 TCAAGCTGACTTCAAGTCCCAAGTCTACTACTTCAAGGCGGAAAGCGAGTACAGCTTGC 840
QY 3053 AAAGGTGATGGAATGATCGCAGTGCACCAAGCTTGCCTGCGGACCCACGTTTGA 3112
DB 841 AAAGGTGATGGAATGATCGCAGTGCACCAAGCTTGCCTGCGGACCCACATCTTGA 900
QY 3113 GCCAAGAAAGTCTTGTGTA 3134
DB 901 GTCAAGAAAGTCTTGTGTA 922

Search completed: May 22, 2006, 07:59:20
Job time : 11576.6 secs

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 22, 2006, 00:29:52 ; Search time 1226.93 Seconds
(without alignments)
17815.260 Million cell updates/sec

Title: US-09-555-342b-1_COPY_49_3183

Perfect score: 3135
Sequence: 1 atcgaggaatcagacagcagag.....acaaagagctctctgtgtrac 3135

Scoring table: IDENTITY_NUC
Gapox 10.0 , Gapext 1.0

Searched: 5244920 seqs, 3486124231 residues

Total number of hits satisfying chosen parameters: 10489840

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: N_Geneseq_8: *
2: geneseqn1980s: *
3: geneseqn1980s: *
4: geneseqn2000s: *
5: geneseqn2001as: *
6: geneseqn2001bs: *
7: geneseqn2002as: *
8: geneseqn2002bs: *
9: geneseqn2003as: *
10: geneseqn2003bs: *
11: geneseqn2003as: *
12: geneseqn2004as: *
13: geneseqn2004bs: *
14: geneseqn2005s: *
15: geneseqn2006s: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3135	100.0	3442	2	AAX79183 Human cho
2	3135	100.0	3442	15	AEB03370 Acute mye
3	3135	99.9	5028	14	AEB22714 Human col
4	3131.8	99.9	4687	3	AAAO8582 Human cyt
5	3119.2	99.5	3187	5	AAS64760 DNA encod
6	1431.2	45.7	3094	3	AAC88992 Human pan
7	1037.4	33.1	3997	12	ADQ84807 Human tum
8	1037.4	33.1	3997	13	ADQ84807 Human tum
9	1037.4	33.1	3997	13	ADQ84807 Human tum
10	550.4	17.6	582	10	ACD92278 Human col
11	550.4	17.6	582	10	ACD92278 Human col
12	447	14.3	3200	12	ADQ84832 Human col
13	420.8	13.4	435	14	ACU61462 Human col
14	403.4	12.9	637	4	AA627016 Human col
15	402.4	12.8	1718	4	AA627016 Human col
16	402.4	12.8	1718	10	ADQ84832 Human col
17	385.6	12.3	393	9	ACH21582 Human adu
18	358.4	11.4	506	6	ABL67912 Ovary can

19	356.8	11.4	671	13	ADQ52709	Adg52709 Novel can
20	334	10.7	441	15	AEB88734	Aef88734 Human CDN
21	306.6	9.8	3275	13	ADP07347	Adp07347 Full leng
22	304.8	9.7	2277	10	ADCT1321	Adct1321 Human col
23	297	9.5	379	8	ABX34989	Abx34989 Bovine ES
24	292	9.3	400	8	ABX34991	Abx34991 Bovine ES
25	254	8.1	1824	5	AAS64761	Aas64761 DNA encod
26	233.4	7.4	6503	13	ACN41157	Acn41157 Human dia
27	229	7.3	1220	3	ABL02335	Ab102335 Drosophi1
28	223.6	7.1	241	4	AAA42150	Aaa42150 Human sec
29	207.8	6.6	716	4	AA627440	Aa627440 cDNA enco
30	207.8	6.6	716	10	ADB93618	Adb93618 Human CDN
31	202.8	6.5	1747	6	ABL03867	Ab103867 Drosophi1
32	200.6	6.4	717	6	ACB08828	Ab190828 Human pol
33	192.4	6.1	2410	8	ACC44336	Acc44336 Gene enco
34	192.4	6.1	2636	14	AEA19659	Aea19659 Novel hum
35	192.4	6.1	2867	12	ADP10335	Adp10335 Reference
36	192.4	6.1	2867	13	ADR90365	Adr90365 Human ful
37	192.4	6.1	2867	13	ADP54688	Adp54688 Human pro
38	192.4	6.1	2867	14	ADY15865	Ady15865 DNA encod
39	192.4	6.1	4338	14	AEA19657	Aea19657 Novel hum
40	192.4	6.1	4922	5	ABV30214	Abv30214 Human pro
41	187	6.0	2882	14	AD62844	Ad62844 Murine pr
42	181.8	5.8	2781	14	ADY43235	Ady43235 Human psy
43	181.8	5.8	3643	6	ABK84074	Abk84074 Human CDN
44	181.8	5.8	3643	13	ADR40158	Adr40158 Human pro
45	181.8	5.8	3963	13	ADP54549	Adp54549 Human pro

ALIGNMENTS

RESULT 1	
AAX79183	
ID	AAX79183 standard; DNA; 3442 BP.
XX	
XX	AAX79183;
AC	
XX	
DT	17-AUG-1999 (first entry)
XX	
DE	Human chondrocyte-derived gene CDEP.
XX	
KW	Differentiation; human; foetal; chondrocyte; ezrin-like domain; cancer;
KM	DB1 homology domain; pleckstrin homology domain; rheumatoid 'arthritis;
KM	drug; ss.
XX	
OS	Homo sapiens.
XX	
PN	MO928458-A1.
XX	
PD	10-JUN-1999.
XX	
PF	27-NOV-1998; 98WO-JP005348.
XX	
PR	27-NOV-1997; 97JP-00342060.
XX	
PA	(CHUS) CHUGAI SEIYAKU KK.
XX	
PI	Kato Y, Kawamoto T, Koyano Y;
XX	
DR	WPI; 1999-371117/31.
XX	
DR	P-PSDB; AAY07482.
PT	Protein CDEP expressed in differentiated chondrocytes, and gene encoding
XX	it.
XX	
PS	Claim 5; Fig 1; 59pp; Japanese.
XX	
CC	This sequence represents the coding region for a protein (CDEP) expressed
CC	in differentiated human foetal chondrocytes, which contains an ezrin-like
CC	domain, a DB1 homology (DH) domain and a pleckstrin homology (PH) domain.
CC	The nucleic acid or protein can be used in the investigation and
CC	treatment of cancers and arthritic diseases (including chronic rheumatoid

CC arthritis), or for screening of candidate anticancer drugs
XX Sequence 3442 BP; 864 A; 952 C; 927 G; 699 T; 0 U; 0 Other;

Query Match 100.0%; Score 3135; DB 2; Length 3442;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ATGGGGAATGATGAGAGAGGCGCCACCGAGATCGACTGGGGGCCCCGGAAAAATTCG 60
DB |||||
QY 49 ATGGGGAATGATGAGAGAGGCGCCACCGAGATCGACTGGGGGCCCCGGAAAAATTCG 108
DB |||||
QY 61 GGGATCAGTACCTTGGAAAGTGGACAGAAAGCCGCCCAACACTTTCAGAAAACTCGTG 120
DB |||||
QY 109 GGGATCAGTACCTTGGAAAGTGGACAGAAAGCCGCCCAACACTTTCAGAAAACTCGTG 168
DB |||||
QY 121 TCCATCAAAATCCAGATGCTGGATGACCCAGAGAGCATTTGAAGTTCCACAAAGAGCT 180
DB |||||
QY 169 TCCATCAAAATCCAGATGCTGGATGACCCAGAGAGCATTTGAAGTTCCACAAAGAGCT 228
DB |||||
QY 181 CCTGGGAAGGTGCTGCTGATGATGATTCGAAACCTCAACCTCGTGGAAAGTGACTAT 240
DB |||||
QY 229 CTTGGGAAGGTGCTGCTGATGATGATTCGAAACCTCAACCTCGTGGAAAGTGACTAT 288
DB |||||
QY 241 TTTGGCTCGAGTTTCTGATCACAAAAAGATCAGGTGTGCTGATCTCTTAAACC 300
DB |||||
QY 289 TTTGGCTCGAGTTTCTGATCACAAAAAGATCAGGTGTGCTGATCTCTTAAACC 348
DB |||||
QY 301 ATTGTGAAACAGATTGAAAGGCCAAGCAGTGTGTTTAAGTTTGTGTAATTTCTTT 360
DB |||||
QY 349 ATTGTGAAACAGATTGAAAGGCCAAGCAGTGTGTTTAAGTTTGTGTAATTTCTTT 408
DB |||||
QY 361 CGGCTGACACACACAACTCCAAAGAAAGTCACAAGGTAACCTGTTCCGCTGACAGTG 420
DB |||||
QY 409 CGGCTGACACACACAACTCCAAAGAAAGTCACAAGGTAACCTGTTCCGCTGACAGTG 468
DB |||||
QY 421 AAGCAGAGTCTGGCTCAAGCAGGTGACGTGTATGACACAGCGCAGCTCTTGAAT 480
DB |||||
QY 469 AAGCAGAGTCTGGCTCAAGCAGGTGACGTGTATGACACAGCGCAGCTCTTGAAT 528
DB |||||
QY 481 TCACACATTGTGCAATCTGAGATTGGGGATTTTGAAGCCTTGGACAGAGAGCATTA 540
DB |||||
QY 529 TCACACATTGTGCAATCTGAGATTGGGGATTTTGAAGCCTTGGACAGAGAGCATTA 588
DB |||||
QY 541 GCAAAAATTAATATACATCTCAGAAAGCAGCTAGAGGACAAATTCGTGAATTTTCA 600
DB |||||
QY 589 GCAAAAATTAATATACATCTCAGAAAGCAGCTAGAGGACAAATTCGTGAATTTTCA 648
DB |||||
QY 601 CATTAACCATTTGACAAACACCAAGAGATCAGATTTTCAAGCTCTAGAGATTGCCGT 660
DB |||||
QY 649 CATTAACCATTTGACAAACACCAAGAGATCAGATTTTCAAGCTCTAGAGATTGCCGT 708
DB |||||
QY 661 CGGCTAGAGATGTAATGGAATCCGGTTGGACCCGGCCAAAGACAGAGGAAGGACAGAAATC 720
DB |||||
QY 709 CGGCTAGAGATGTAATGGAATCCGGTTGGACCCGGCCAAAGACAGAGGAAGGACAGAAATC 768
DB |||||
QY 721 AATCTGGCCGTTGGCCAAACAGGGAATTTCTAGTGTTCAGGGTTTCACTAAGATCAATGCC 780
DB |||||
QY 769 AATCTGGCCGTTGGCCAAACAGGGAATTTCTAGTGTTCAGGGTTTCACTAAGATCAATGCC 828
DB |||||
QY 781 TTCAACTGGGCCCAAGGTGCGGAAAGCTGAGCTTCAAGAGAAAGCGCTTTCATCAAGCTC 840
DB |||||
QY 829 TTCAACTGGGCCCAAGGTGCGGAAAGCTGAGCTTCAAGAGAAAGCGCTTTCATCAAGCTC 888
DB |||||
QY 841 CGGCCGAGTCCCAATAGTCCGTAACAGAGTACCTTGGAAATTCCTGATGGCCAGTCGGGAT 900
DB |||||
QY 889 CGGCCGAGTCCCAATAGTCCGTAACAGAGTACCTTGGAAATTCCTGATGGCCAGTCGGGAT 948
DB |||||
QY 901 TTCTGCAAGTCTTTCGAAAAATCTGTGTGAACATCATGCTTCTTGAAGCTTTTGA 960
DB |||||
QY 949 TTCTGCAAGTCTTTCGAAAAATCTGTGTGAACATCATGCTTCTTGAAGCTTTTGA 1008
DB |||||
QY 961 GAGCCCAAAACAAAGCCCAAGCCGTCCTTTAGCCGGGGGTCAATCATTTGGTTCAAGT 1020
DB |||||
```

```
DB |||||
QY 1009 GAGCCCAAAACAAAGCCCAAGCCGTCCTTTAGCCGGGGGTCAATCATTTGGTTCAAGT 1068
DB |||||
QY 1021 GGTGGGACTCAGAAACAGGTTTTCAGATATGTTAAAGAGAGCATTAAGAAAGGTGACG 1080
DB |||||
QY 1069 GGTGGGACTCAGAAACAGGTTTTCAGATATGTTAAAGAGAGCATTAAGAAAGGTGACG 1128
DB |||||
QY 1081 TTTGAAAGGAAGCAGACAGCAAGATTCATCTATCCGAGCCTTGCTTCACAGCCTACAGA 1140
DB |||||
QY 1129 TTTGAAAGGAAGCAGACAGCAAGATTCATCTATCCGAGCCTTGCTTCACAGCCTACAGA 1188
DB |||||
QY 1141 CTGAATTGGGAAGTGTGGAGCATCTCAGACAGCAACAGCTTACATTTGAGAGAGT 1200
DB |||||
QY 1189 CTGAATTGGGAAGTGTGGAGCATCTCAGACAGCAACAGCTTACATTTGAGAGAGT 1248
DB |||||
QY 1201 GCGGAATCTCCAGGGGGCCAGAGCTGCGGGGAGGAAAGGAAAGGTTTCCGCGGG 1260
DB |||||
QY 1249 GCGGAATCTCCAGGGGGCCAGAGCTGCGGGGAGGAAAGGAAAGGTTTCCGCGGG 1308
DB |||||
QY 1261 GAGCCGGGGTCCACCCGAGCCCTGCGCGAGAGAAAGCCCGGGTTAAACAAGAGGCG 1320
DB |||||
QY 1309 GAGCCGGGGTCCACCCGAGCCCTGCGCGAGAGAAAGCCCGGGTTAAACAAGAGGCG 1368
DB |||||
QY 1321 GACGGAAGCCGCTCGGCGCCACGGAAGAAAGAGAGAGTGTTAAGATATGACCCAG 1380
DB |||||
QY 1369 GACGGAAGCCGCTCGGCGCCACGGAAGAAAGAGAGTGTTAAGATATGACCCAG 1428
DB |||||
QY 1381 CAGAGTAAACCTCAGCCCCGAGCCCAAGCAGAGCTCCCTGACTGGCAGTCTCACTT 1440
DB |||||
QY 1429 CAGAGTAAACCTCAGCCCCGAGCCCAAGCAGAGCTCCCTGACTGGCAGTCTCACTT 1488
DB |||||
QY 1441 TCCGAGTGTCTGTGAATCTCGAGGGGGAGTGGCCCTGCCAAGTGAACCTTGTCTCC 1500
DB |||||
QY 1489 TCCGAGTGTCTGTGAATCTCGAGGGGGAGTGGCCCTGCCAAGTGAACCTTGTCTCC 1548
DB |||||
QY 1501 AACCTGAAGCCCGGACACCAAGAGGCTCTCCCTTGAATCAGCCGCTGTGAATGACAG 1560
DB |||||
QY 1549 AACCTGAAGCCCGGACACCAAGAGGCTCTCCCTTGAATCAGCCGCTGTGAATGACAG 1608
DB |||||
QY 1561 GCTGCCCCCGGACGGAAGATGAGAGTGGGGCCGGAAGGAAGATTCCCACTGATAA 1620
DB |||||
QY 1609 GCTGCCCCCGGACGGAAGATGAGAGTGGGGCCGGAAGGAAGATTCCCACTGATAA 1668
DB |||||
QY 1621 GCGTACTTCAATAGCTAAGAAAGTGTCTACCAAGAGCAATATCTGAAGATCTCGAA 1680
DB |||||
QY 1669 GCGTACTTCAATAGCTAAGAAAGTGTCTACCAAGAGCAATATCTGAAGATCTCGAA 1728
DB |||||
QY 1681 GTTATCACTTGTGTTTCAAGCAAGTGAAGAAAGAGAGCCATGCGGAGAGCATG 1740
DB |||||
QY 1729 GTTATCACTTGTGTTTCAAGCAAGTGAAGAAAGAGAGCCATGCGGAGAGCATG 1788
DB |||||
QY 1741 AAAAGTCAATATTCGGAATTTTGAACCTTGGCAAAATTTCTACTAATTTTCTCAAG 1800
DB |||||
QY 1789 AAAAGTCAATATTCGGAATTTTGAACCTTGGCAAAATTTCTACTAATTTTCTCAAG 1848
DB |||||
QY 1801 GAAATTGAGCAACGACTTGCCTGTGGGAAGGCGCTCAATGCCCAATTCAGAGATTAC 1860
DB |||||
QY 1849 GAAATTGAGCAACGACTTGCCTGTGGGAAGGCGCTCAATGCCCAATTCAGAGATTAC 1908
DB |||||
QY 1861 CAAAGATTCGGGAGTGTATGTGAAGAACTTTCAGGGCATTAAGCATCTGGCGGCTCAG 1920
DB |||||
QY 1909 CAAAGATTCGGGAGTGTATGTGAAGAACTTTCAGGGCATTAAGCATCTGGCGGCTCAG 1968
DB |||||
QY 1921 CTGTGGAAGCAGAGGAGCTTGGAGGCGCTGGAAGATGGAATCAAGAGCTCCGCGCG 1980
DB |||||
QY 1969 CTGTGGAAGCAGAGGAGCTTGGAGGCGCTGGAAGATGGAATCAAGAGCTCCGCGCG 2028
DB |||||
QY 1981 CTGGAAGACTTCTGCAAGACTTTGAGCTGCAAGAGTGTGTACTTCAACGCTCAACACC 2040
DB |||||
QY 2029 CTGGAAGACTTCTGCAAGACTTTGAGCTGCAAGAGTGTGTACTTCAACGCTCAACACC 2088
DB |||||
QY 2041 TTCTCTCGGGCCACTGACCGGCTCATGCACTAACAGAGTCTTGAAGGCGGCTGTGC 2100
DB |||||
```

```

Db      2089 TTCTCTGGGCGCACTGCAACGGCTCATGACATGACAGAGGTCCTGAGCGGCTGTGC 2148
Qy      2101 AAMCACCACCCCGGAGCCAGCCGATTTCAAGGACTGCGGAGCCGCTTTGGCAGAGATC 2160
Db      2149 AAMCACCACCCCGGAGCCAGCCGATTTCAAGGACTGCGGAGCCGCTTTGGCAGAGATC 2208
Qy      2161 ACGGAGATGATGACAGCTCCAGCGATGATGATCAAGATGAGAAATTTCCAGAAAGCTG 2220
Db      2209 ACGGAGATGATGACAGCTCCAGCGATGATGATCAAGATGAGAAATTTCCAGAAAGCTG 2268
Qy      2221 CACGAATCAAGAAAGATTTGATTTGAGATGACAAATCTTGTGTGTTCCGGGAGAGAGTTTC 2280
Db      2269 CACGAATCAAGAAAGATTTGATTTGAGATGACAAATCTTGTGTGTTCCGGGAGAGAGTTTC 2328
Qy      2281 ATCCGCTGTGGGAGCCCTCAGCAAGCTTCCGGGAGAGAGAGGCTCCAGCAGCCGATGTTCTTC 2340
Db      2329 ATCCGCTGTGGGAGCCCTCAGCAAGCTTCCGGGAGAGAGAGGCTCCAGCAGCCGATGTTCTTC 2388
Qy      2341 CTGTTCAACGACGTCCTGCTATACAGAGCCGGGGGCTGACGGGCTCCAAATCAGTTTAA 2400
Db      2389 CTGTTCAACGACGTCCTGCTATACAGAGCCGGGGGCTGACGGGCTCCAAATCAGTTTAA 2448
Qy      2401 GTCCACGGGAGCTCCCGCTCTATGCGATGACGATTGAGGAGACGAAAGCAGTGGGG 2460
Db      2449 GTCCACGGGAGCTCCCGCTCTATGCGATGACGATTGAGGAGACGAAAGCAGTGGGG 2508
Qy      2461 GTGCCCCACTGCTGACCTTCGCGGGCCAGCGGCACTCATCTGTTGGCCGCAATTTCT 2520
Db      2509 GTGCCCCACTGCTGACCTTCGCGGGCCAGCGGCACTCATCTGTTGGCCGCAATTTCT 2568
Qy      2521 CGGTCGGAATGAGAAAGTGGTTGAGGACATCGAGTGGCCATTGACCTGGCGGAGAAAG 2580
Db      2569 CGGTCGGAATGAGAAAGTGGTTGAGGACATCGAGTGGCCATTGACCTGGCGGAGAAAG 2628
Qy      2581 AGCAGCAGCCCGCCCTGAGTTCTTGCCAGCAGCCCGCTGACAAAGTCCCTGAT 2640
Db      2629 AGCAGCAGCCCGCCCTGAGTTCTTGCCAGCAGCCCGCTGACAAAGTCCCTGAT 2688
Qy      2641 GAAGCCACCGGCTGACCGAGATCAGAGATGACCTGAGCCCTTCGCGCAATGCTG 2700
Db      2689 GAAGCCACCGGCTGACCGAGATCAGAGATGACCTGAGCCCTTCGCGCAATGCTG 2748
Qy      2701 GAGCGCAGGCGCCCGCAGCCGCGCAACAAATGTTGACGTGTGCGCACCGCAACACC 2760
Db      2749 GAGCGCAGGCGCCCGCAGCCGCGCAACAAATGTTGACGTGTGCGCACCGCAACACC 2808
Qy      2761 AGCGTCTCCATGATGATTCAGCATCGCAGTGAAGATCAGTTGTCTGGAACCTGCTG 2820
Db      2809 AGCGTCTCCATGATGATTCAGCATCGCAGTGAAGATCAGTTGTCTGGAACCTGCTG 2868
Qy      2821 AGGAATTTCAAAACAGCAACGGGTGCGAAGCTGTGGGTGTTTCAAACTTCTGC 2880
Db      2869 AGGAATTTCAAAACAGCAACGGGTGCGAAGCTGTGGGTGTTTCAAACTTCTGC 2928
Qy      2881 CTGTTCTTCAAAATCACACAGCAATCATCCCTTGCAGGCTGCTGCTCGGC 2940
Db      2929 CTGTTCTTCAAAATCACACAGCAATCATCCCTTGCAGGCTGCTGCTCGGC 2988
Qy      2941 TACTGCTCAACATCCCTCTGATGCTCGAAGACATCCAGGCTGCTGCTCGGC 3000
Db      2989 TACTGCTCAACATCCCTCTGATGCTCGAAGACATCCAGGCTGCTGCTCGGC 3048
Qy      3001 CACTTCAAGTCCCAAGTCTACTTCTCAAGGCGGAAAGCAAGTACAGTTTCAAAAGCTG 3060
Db      3049 CACTTCAAGTCCCAAGTCTACTTCTCAAGGCGGAAAGCAAGTACAGTTTCAAAAGCTG 3108
Qy      3061 ATGAAGATGATCCGAGTGCACCAAGCTGCTGCTCGGAGCCCAAGTGTGAGCAAAA 3120
Db      3109 ATGAAGATGATCCGAGTGCACCAAGCTGCTGCTCGGAGCCCAAGTGTGAGCAAAA 3168
Qy      3121 GAGTCTCTGTGTAT 3135
Db      3169 GAGTCTCTGTGTAT 3183

```

```

RESULT 2
AE03370
ID ABE03370 standard; DNA; 3442 BP.
XX
AC ABE03370;
XX
DT 09-FEB-2006 (first entry)
XX
DE Acute myelogenous leukemia prognosis related DNA sequence SEQ ID NO: 31.
XX
KW ds: gene; acute myelogenous leukemia; prognosis; gene expression;
KW bioclip.
XX
OS Homo sapiens.
XX
PN JF200533987-A.
XX
PD 08-DEC-2005.
XX
PF 06-MAY-2005; 2005JP-00135284.
XX
PR 06-MAY-2004; 2004US-0568635P.
XX
PA (VERI-) VERIDEX LLC.
XX
PI Laponi M.
XX
DR WPI; 2006-004067/01.
XX
PT Evaluating an acute myelogenous leukemia patient's prognosis, comprising
PT detecting lower or higher expression level of gene recognized by probe
PT set e.g. 202820-at and 206148-at, with respect to predetermined cut-off
PT level.
XX
PS Example 5; SEQ ID NO 31; 60pp; Japanese.
XX
XX
CC The present sequence is that of a human gene sequence which is claimed
CC for use in evaluating the prognosis of patients suffering from acute
CC myelogenous leukemia (AML) by analysis of the patient's gene expression
CC profile. The invention relates to a novel method for evaluating the
CC prognosis of a patient with acute myelogenous leukemia by detecting a
CC higher/lower expression level of genes encoding mRNA recognized by a
CC probe set chosen from 19 nucleotide sequences (ABE03340 or ABE03344-
CC ABE03361). Also, gene expression profiling of AML patients using a set of
CC probes based on 167 sequences (ABE03362-ABE03527) using a bioclip was
CC performed to determine prognosis. Methods are also included for
CC determining an AML patient's state, determining AML patient's treatment,
CC protocol, determining whether the patient will respond to the treatment,
CC and producing an AML patient's prognosis report by analysis of the
CC patient's gene expression profile. The methods are useful in evaluating an
CC AML patient's prognosis and for treating AML patient.
XX
SQ Sequence 3442 BP; 864 A; 952 C; 927 G; 699 T; 0 U; 0 Other;

```

Query Match 100.0%; Score 3135; DB 15; Length 3442;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy      1 ATGGGAGAAATGAGCAGAGCGCGACCCGAGATCAGACTGGGGCCCCGGGAAATTCG 60
Db      49 ATGGGAGAAATGAGCAGAGCGCGACCCGAGATCAGACTGGGGCCCCGGGAAATTCG 108
Qy      61 GGGATCAGTACTTGGAAAGTGAACAAGACCCGCCCAACCTTCAGAGAAACTCGTG 120
Db      109 GGGATCAGTACTTGGAAAGTGAACAAGACCCGCCCAACCTTCAGAGAAACTCGTG 168
Qy      121 TCCATCAAAATTCAGATGTGATGACACCCAGAGGCACTTTGAAGTTCCAAAGAGCT 180
Db      169 TCCATCAAAATTCAGATGTGATGACACCCAGAGGCACTTTGAAGTTCCAAAGAGCT 228
Qy      181 CTTGGGAAGTGTGCTGATGACATTTGCAACCACTCAACCTGTGGAAGTGAATAT 240

```

Db 229 CCGGGAAGTGTGCTGATGATGCACTTGCACCACTCGTGAAGGTGACTAT 288
Qy 241 TTTGGCTCGAGTTTCTCTATCAAAAAGATCAAGGTGTGGATCTCTTAAACC 300
Db 289 TTTGGCTCGAGTTTCTCTATCAAAAAGATCAAGGTGTGGATCTCTTAAACC 348
Qy 301 ATTGTAAACAGATTGAAGGCCAAGCAAGTTGTGTTAAGTTTGGTAAATCTTT 360
Db 349 ATTGTAAACAGATTGAAGGCCAAGCAAGTTGTGTTAAGTTTGGTAAATCTTT 408
Qy 361 CCGCTGACCAACACAACTCCAAAGAACTTCACAAAGTACTGTTCGCTGCAGTG 420
Db 409 CCGCTGACCAACACAACTCCAAAGAACTTCACAAAGTACTGTTCGCTGCAGTG 468
Qy 421 AAGCAGGACTTGGCTCAAGGCAAGTTGAGTGTAAAGACACAGCCGACTCTTGAT 480
Db 469 AAGCAGGACTTGGCTCAAGGCAAGTTGAGTGTAAAGACACAGCCGACTCTTGAT 528
Qy 481 TCACACATTGTGCAATCTGAGATTGGGGAATTTGATGAAGCTTGGACAGAGCACTTA 540
Db 529 TCACACATTGTGCAATCTGAGATTGGGGAATTTGATGAAGCTTGGACAGAGCACTTA 588
Qy 541 GCAAAAATTAATACATACCTCAGCAAGACGCACTAGAGCAAAATCGTGAATTTCA 600
Db 589 GCAAAAATTAATACATACCTCAGCAAGACGCACTAGAGCAAAATCGTGAATTTCA 648
Qy 601 CATTAACCAATTGGAACAAACACAGAGAAATGATTTCCAGTCCCTAAGATTCGCGT 660
Db 649 CATTAACCAATTGGAACAAACACAGAGAAATGATTTCCAGTCCCTAAGATTCGCGT 708
Qy 661 CCGCTAGAGATGATGAATCCGGTTGCAACCGGCCAAGGACAGGAGGCAAGAGATC 720
Db 709 CCGCTAGAGATGATGAATCCGGTTGCAACCGGCCAAGGACAGGAGGCAAGAGATC 768
Qy 721 AATCTGCGGTGCGCAACCGGAAATCTAGTGTTCAGGTTTCACTAAGATCAATGCG 780
Db 769 AATCTGCGGTGCGCAACCGGAAATCTAGTGTTCAGGTTTCACTAAGATCAATGCG 828
Qy 781 TTCAACTGGGCCAAGGTGGGAAAGCTGAGCTTCAAGAGAAAGGCTTTCATCAAGCTC 840
Db 829 TTCAACTGGGCCAAGGTGGGAAAGCTGAGCTTCAAGAGAAAGGCTTTCATCAAGCTC 888
Qy 841 CCGCCAGATGCCAATAGTCGTACAGAGATACCTTGAATTCCTGATGGCAATCGGAT 900
Db 889 CCGCCAGATGCCAATAGTCGTACAGAGATACCTTGAATTCCTGATGGCAATCGGAT 948
Qy 901 TTCTGCAAGTCTTCTGAAAATCTGTGTGAACATCATGCTTCTTTAGACTTTTGA 960
Db 949 TTCTGCAAGTCTTCTGAAAATCTGTGTGAACATCATGCTTCTTTAGACTTTTGA 1008
Qy 961 GAGCCCAAAACAAAGGCCAAGCCGCTCTTTAGCCGGGGGTCATCATTTGGGTTCA 1020
Db 1009 GAGCCCAAAACAAAGGCCAAGCCGCTCTTTAGCCGGGGGTCATCATTTGGGTTCA 1068
Qy 1021 GGTTCGAGCTCAGAGAGGTTCTGACTATGTTAAAGAGAGACATAAGAGGTGAG 1080
Db 1069 GGTTCGAGCTCAGAGAGGTTCTGACTATGTTAAAGAGAGACATAAGAGGTGAG 1128
Qy 1081 TTTGAAAGAGAGACAGCAAGATTCAATTATCCGAGCCCTTGCCTTCAAGCTTACAG 1140
Db 1129 TTTGAAAGAGAGACAGCAAGATTCAATTATCCGAGCCCTTGCCTTCAAGCTTACAG 1188
Qy 1141 CTGAATTCCGAAGTGTGAGAGCTCAGAGAGACAGACCTTTCATTTGGAGAAAGGT 1200
Db 1189 CTGAATTCCGAAGTGTGAGAGCTCAGAGAGACAGACCTTTCATTTGGAGAAAGGT 1248
Qy 1201 GCGGAATCTCAGAGGGGCGAGAGCTCGCGAGAGAAAGAAACGAAGGTTTCCGCGGG 1260
Db 1249 GCGGAATCTCAGAGGGGCGAGAGCTCGCGAGAGAAAGAAACGAAGGTTTCCGCGGG 1308
Qy 1261 GAGCCGGGGTCCGACCCGAGCCCTGCGCGAGAGAAAGCCCGCGGGTTAACAGCGGGG 1320

Db 1309 GAGCCGGGGTCCGACCCGAGCCCTGCGCGAGAGAAAGCCCGCGGGTTAACAGCGGGG 1368
Qy 1321 GAGCGAGCGGCTCGGCGCCCAACGAGAGAAAGAGAGAGGTGTTAAGGATAGAGCCGAG 1380
Db 1369 GAGCGAGCGGCTCGGCGCCCAACGAGAGAAAGAGAGAGGTGTTAAGGATAGAGCCGAG 1428
Qy 1381 CAGAGTAAACCTCAGCCCCCGGAGCCAAAGCAAGGCTCCCTGATCTGGAGTCCGACTT 1440
Db 1429 CAGAGTAAACCTCAGCCCCCGGAGCCAAAGCAAGGCTCCCTGATCTGGAGTCCGACTT 1488
Qy 1441 TCCGAGCTGTCTGTAACTCCGAGGGGGAGTGGCCCTGCCCAAGTACCTTGTCTCC 1500
Db 1489 TCCGAGCTGTCTGTAACTCCGAGGGGGAGTGGCCCTGCCCAAGTACCTTGTCTCC 1548
Qy 1501 AACCTGAGCCCGGACCAAGAGGCTCTCCCTTATCAGCCGCTGCTGAATGACAG 1560
Db 1549 AACCTGAGCCCGGACCAAGAGGCTCTCCCTTATCAGCCGCTGCTGAATGACAG 1608
Qy 1561 GCTTCCCGGACGAGAGATGAGGATGAGGGCCGAGAGAAAGATTCCCACTGATAA 1620
Db 1609 GCTTCCCGGACGAGAGATGAGGATGAGGGCCGAGAGAAAGATTCCCACTGATAA 1668
Qy 1621 GGTACTTCAATAGCTAAGAAAGTCTACACCGAGCCGACATATCTGAAGATCTCGAA 1680
Db 1669 GGTACTTCAATAGCTAAGAAAGTCTACACCGAGCCGACATATCTGAAGATCTCGAA 1728
Qy 1681 GTTATCACTTGGGTTTCAAGACAGTGAAGCAAAAGAGAGGCGCATCGGAGCACTG 1740
Db 1729 GTTATCACTTGGGTTTCAAGACAGTGAAGCAAAAGAGAGGCGCATCGGAGCACTG 1788
Qy 1741 AAAAGTCTATATTCGGAATTTTGAACCTTTGCAAAATTTCAATACTAATTTTCTCAG 1800
Db 1789 AAAAGTCTATATTCGGAATTTTGAACCTTTGCAAAATTTCAATACTAATTTTCTCAG 1848
Qy 1801 GAAATTTGAGCAACGCTTGGCCCTGAGGAAAGGCGGCTCAATGCCAATTCAGAAATTC 1860
Db 1849 GAAATTTGAGCAACGCTTGGCCCTGAGGAAAGGCGGCTCAATGCCAATTCAGAAATTC 1908
Qy 1861 CAAGAATCGGAGATGATGCTGAAGAAATTCAGGGCAATGAAGCACTTGGCGGCTCAG 1920
Db 1909 CAAGAATCGGAGATGATGCTGAAGAAATTCAGGGCAATGAAGCACTTGGCGGCTCAG 1968
Qy 1921 CTGTGAAAGCACAGGAGGCTTTGAGGCGCTGAGAAATGAATCAAGACTCCGCGCG 1980
Db 1969 CTGTGAAAGCACAGGAGGCTTTGAGGCGCTGAGAAATGAATCAAGACTCCGCGCG 2028
Qy 1981 CTGAGAACTTCTGAGAGACTTTGAGCTGAGAAAGGTGTGTTAATCTTCCGCTCAAC 2040
Db 2029 CTGAGAACTTCTGAGAGACTTTGAGCTGAGAAAGGTGTGTTAATCTTCCGCTCAAC 2088
Qy 2041 TTCTCTGCGGCACTGCAACCGGCTCATGCACTTCAAGCAGAGTCTGAGAGCGGCTGTC 2100
Db 2089 TTCTCTGCGGCACTGCAACCGGCTCATGCACTTCAAGCAGAGTCTGAGAGCGGCTGTC 2148
Qy 2101 AAAACACACCGGCGAGGCAAGCGCTTCAAGGAGCTCCGAGCGCTTTGGCAGAGATC 2160
Db 2149 AAAACACACCGGCGAGGCAAGCGCTTCAAGGAGCTCCGAGCGCTTTGGCAGAGATC 2208
Qy 2161 ACGAGAGATGGGCAACAGCTTCAAGGTAAGATGAACAAATGAGAAATTTCCAGAAAGCTG 2220
Db 2209 ACGAGAGATGGGCAACAGCTTCAAGGTAAGATGAACAAATGAGAAATTTCCAGAAAGCTG 2268
Qy 2221 CAGGAATCAAGAAAGATTGATTTGGATTGCAATCTTGTGTGTTCCGGAGAGGAGTTT 2280
Db 2269 CAGGAATCAAGAAAGATTGATTTGGATTGCAATCTTGTGTGTTCCGGAGAGGAGTTT 2328
Qy 2281 ATCGGTCTGGGAGCTTCAAGAGCTTCTCGGGAAAGGAGCTTCAAGAGCGCATGTTCTT 2340
Db 2329 ATCGGTCTGGGAGCTTCAAGAGCTTCTCGGGAAAGGAGCTTCAAGAGCGCATGTTCTT 2388
Qy 2341 CTGTTCAACAGCTCTCTGTTATACAGAGCGGGGGCTGAGGGCTCCATCATGTTTAA 2400
Db 2389 CTGTTCAACAGCTCTCTGTTATACAGAGCGGGGGCTGAGGGCTCCATCATGTTTAA 2448

QY 2401 GTCCAGGGAGGCTCCCGCTCTTAATGCAATGAGAGAGCCAAAGCAAGTGGGG 2460
 DB 2449 GTCCAGGGAGGCTCCCGCTCTTAATGCAATGAGAGAGCCAAAGCAAGTGGGG 2508
 QY 2461 GTGCCCCACTGCTGACCTCCGCGGCGCAGGCAAGTCCATCATGCGGCGCAGTTC 2520
 DB 2509 GTGCCCCACTGCTGACCTCCGCGGCGCAGGCAAGTCCATCATGCGGCGCAGTTC 2568
 QY 2521 CGGTCCGAGATGAGAAATGGGTTGAGAGCATTCAGATGGCCATTGACTGGCGAGAA 2580
 DB 2569 CGGTCCGAGATGAGAAATGGGTTGAGAGCATTCAGATGGCCATTGACTGGCGAGAA 2628
 QY 2581 AGCAGCAGCCCCCGCTGAGTTCTCTGCGCAGCAGCCCCCTTCAACAAGTCCCTGAT 2640
 DB 2629 AGCAGCAGCCCCCGCTGAGTTCTCTGCGCAGCAGCCCCCTTCAACAAGTCCCTGAT 2688
 QY 2641 GAAGCCACCGCGCTGACCGAGATGACCTGAGCGCTTCGCGCAATCGCTG 2700
 DB 2689 GAAGCCACCGCGCTGACCGAGATGACCTGAGCGCTTCGCGCAATCGCTG 2748
 QY 2701 GAGCGCAGGCGCCCGCAGCGCGGCAACAAATGCTGCTGCGCAGCGCAACCC 2760
 DB 2749 GAGCGCAGGCGCCCGCAGCGCGGCAACAAATGCTGCTGCGCAGCGCAACCC 2808
 QY 2761 AGGCTCTCCATGGTGAATTCAGATTCGCGAGTGAATGATGTTCTGGAACCTGCTG 2820
 DB 2809 AGGCTCTCCATGGTGAATTCAGATTCGCGAGTGAATGATGTTCTGGAACCTGCTG 2868
 QY 2821 AGGAATTCMAAAGCAGCAACGGGTGGCAGAGCTGAGTGGTGTTCACAACTTCTGC 2880
 DB 2869 AGGAATTCMAAAGCAGCAACGGGTGGCAGAGCTGAGTGGTGTTCACAACTTCTGC 2928
 QY 2881 CTGTTCTTCAAAATCACAACAGCAATATCCCTTGCAGACCTGCTGCTCGGC 2940
 DB 2929 CTGTTCTTCAAAATCACAACAGCAATATCCCTTGCAGACCTGCTGCTCGGC 2988
 QY 2941 TACTCGCTCAACCTCCCTGCTGAGTCCGAGAAATCCAGAAATCAACGTTCAAGCTG 3000
 DB 2989 TACTCGCTCAACCTCCCTGCTGAGTCCGAGAAATCCAGAAATCAACGTTCAAGCTG 3048
 QY 3001 CACTTCAAGTCCAGCTCTACTTCAAGGCGGAAAGCAGATCAAGTTCGAAAGGTG 3060
 DB 3049 CACTTCAAGTCCAGCTCTACTTCAAGGCGGAAAGCAGATCAAGTTCGAAAGGTG 3108
 QY 3109 ATGGAATGATCCGAGTGCACCAAGCTGCGGAGCCCAAGTTCGAGCAACA 3168
 DB 3169 GAGTCTCTGTGTAT 3183
 QY 3183 GAGTCTCTGTGTAT 3183
 DB 3183 GAGTCTCTGTGTAT 3183
 RESULT 3
 AEB22714
 ID AEB22714 standard, DNA, 5028 BP.
 AC AEB22714;
 XX
 DT 22-SEP-2005 (first entry)
 XX
 DE Human colon cancer-associated gene SEQ ID NO:9.
 XX
 DE cancer; genetic marker; lymphoma; cytostatic; hematological disease;
 KM immune disorder; neoplasm; colon tumor; gastrointestinal disease;
 KM adenocarcinoma; carcinoma; screening; diagnostic; prognostic;
 KM pharmaceutical; de; gene.
 XX
 OS Homo sapiens.
 XX
 PN MO2005064009-A1.
 XX

PD 14-JUL-2005.
 XX
 PF 23-DEC-2004; 2004MO-DK000914.
 XX
 PR 27-DEC-2003; 2003DK-00001940.
 PR 24-JAN-2004; 2004DK-0000096.
 PR 07-APR-2004; 2004DK-0000086.
 PR 26-NOV-2004; 2004DK-00001843.
 XX
 PA (AROS-) AROS APPLIED BIOTECHNOLOGY APS.
 PI Oncofct TF, Jensen JL, Kruhooffer M, Laiho P, Aaltonen LA;
 XX
 DR WPI, 2005-506661/51.
 PT Classifying the cancer in an individual having contracted cancer
 PS comprises determining the microsatellite status of the tumor.
 XX
 Claim 9: SEQ ID NO 9; 338bp; English.
 XX
 CC The invention relates to a method of classifying the cancer in an
 CC individual having contracted cancer, which comprises determining the
 CC microsatellite status of the tumor. The microsatellite status or the
 CC hereditary or sporadic nature of the cancer is determined by a method
 CC comprising: (a) obtaining a sample from the individual having contracted
 CC cancer, the sample comprising gene expression products, the presence
 CC and/or amount of which forms a pattern that is indicative of the
 CC microsatellite status or the hereditary or sporadic nature of the cancer;
 CC (b) determining the presence and/or amount of the gene expression
 CC products forming the pattern, and obtaining an indication of the
 CC microsatellite status or the hereditary or sporadic nature of the cancer
 CC in the individual based on step (b). At least one or two of the gene
 CC expression products forming a pattern used to determine the
 CC microsatellite status or the hereditary or sporadic nature of the cancer
 CC are selected individually from any of the 134 genes comprising 367-7850
 CC bp (AEB22706 to AEB22839), given in the specification. Also included are:
 CC a method for treating an individual; a method for reducing malignancy of
 CC a cell; a method for producing antibodies against an expression product
 CC of a cell from a biological tissue; pharmaceutical composition, for
 CC treating a classified cancer, comprising at least one antibody, at least
 CC one polypeptide, or at least one nucleic acid and/or probe; and an assay,
 CC for classification of cancer in an individual having contracted cancer,
 CC comprising at least one marker capable of determining the microsatellite
 CC status in a sample and at least one marker in a sample determining the
 CC prognostic marker, where the microsatellite status and the prognostic
 CC marker is determined simultaneously or sequentially. The method above is
 CC useful for producing an assay for classifying cancer in animal tissue.
 CC The peptide, gene, or probe is useful for preparing a pharmaceutical
 CC composition for the treatment of a cancer in animal tissue. The method is
 CC useful for classifying the cancer in an individual having contracted
 CC cancer, i.e., colon cancer, e.g., an adenocarcinoma, a carcinoma, a
 CC teratoma, a sarcoma, and/or a lymphoma. The present sequence represents a
 CC colon cancer-associated gene used in the method of the invention.
 XX
 SQ Sequence 5028 BP; 1288 A; 1387 C; 1276 G; 1077 T; 0 U; 0 Other;
 Query Match 99.9%; Score 3133.4; DB 14; Length 5028;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 3134; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 ATGGGAGAAATAGAGCAGAGCGGAGCCCAAGATCAAGCTGGGGCCCCGGGAAATTCG 60
 DB 337 ATGGGAGAAATAGAGCAGAGCGGAGCCCAAGATCAAGCTGGGGCCCCGGGAAATTCG 396
 QY 61 GGGATCAGTACCTTGGAACTGTGAGCAGAAAGCCGCCCAACCTTCAGAGAAACTCGTG 120
 DB 397 GGGATCAGTACCTTGGAACTGTGAGCAGAAAGCCGCCCAACCTTCAGAGAAACTCGTG 456
 QY 121 TCCATCAAAATCAGATGCTGTGATGACACCCGAGAGGCTTTGAAAGTTCCAAAGAGCT 180
 DB 457 TCCATCAAAATCAGATGCTGTGATGACACCCGAGAGGCTTTGAAAGTTCCAAAGAGCT 516
 QY 181 CTGGGAGAGTGTGCTGATGCAAGTTTGCAACCACTCAACTCGTGAAGTGAATAT 240

Db 517 CTTGGGAAGTGTCTGTGATGATGCACTTGCACCACTCGTGAAGGTGACTAT 576
Qy 241 TTTGGCTGAGTTTCTCTATCAAAAAGATCAAGGTGTGTGATCTCTTAAACCC 300
Db 577 TTTGGCTGAGTTTCTCTATCAAAAAGATCAAGGTGTGTGATCTCTTAAACCC 636
Qy 301 ATTGTAAACAGATTGAAAGCCAAAGCAAGTGTGTGTTAAGTTTGTGTAATCTTT 360
Db 637 ATTGTAAACAGATTGAAAGCCAAAGCAAGTGTGTGTTAAGTTTGTGTAATCTTT 696
Qy 361 CCGCTGTACACACACAATCTCAAGAAAGACTCAAGGTACTGTTCGCGCTGACAGT 420
Db 697 CCGCTGTACACACACAATCTCAAGAAAGACTCAAGGTACTGTTCGCGCTGACAGT 756
Qy 421 AAGCAGAGACTGTGCTCAAGGAGGTGAGTGTAAAGACCAAGCGCACTCTCTTGAAT 480
Db 757 AAGCAGAGACTGTGCTCAAGGAGGTGAGTGTAAAGACCAAGCGCACTCTCTTGAAT 816
Qy 481 TCACACATTGTGCAATCTGAGATTGGGAAATTTTGAATGAAGCCTTGAAGAGCACTTA 540
Db 817 TCACACATTGTGCAATCTGAGATTGGGAAATTTTGAATGAAGCCTTGAAGAGCACTTA 876
Qy 541 GCAAAAATTAATATACATACCTCTGACAAAGACGCACTAGAGACAAAATCGTGAATTTAC 600
Db 877 GCAAAAATTAATATACATACCTCTGACAAAGACGCACTAGAGACAAAATCGTGAATTTAC 936
Qy 601 CATTAACAATTGGAACAAACACAGAGAAATGCAATTCAGCTCCTAAGATGCGCGT 660
Db 937 CATTAACAATTGGAACAAACACAGAGAAATGCAATTCAGCTCCTAAGATGCGCGT 996
Qy 661 CCGCTAGAGATGATGGAATCCGTTGCAACCGCGCAAGAGACAGAGAGGCAAGAGATC 720
Db 997 CCGCTAGAGATGATGGAATCCGTTGCAACCGCGCAAGAGACAGAGAGGCAAGAGATC 1056
Qy 721 AATCTGCGGTGCAACAGGGAATCTAGTGTTCAGGGTTTCACTAAGATCAATGCTC 780
Db 1057 AATCTGCGGTGCAACAGGGAATCTAGTGTTCAGGGTTTCACTAAGATCAATGCTC 1116
Qy 781 TTCAACTGGGCGAAGGTGGGAAAGCTGAGCTTCAAGAGAGGCTTCTCATCAAGCTC 840
Db 1117 TTCAACTGGGCGAAGGTGGGAAAGCTGAGCTTCAAGAGAGGCTTCTCATCAAGCTC 1176
Qy 841 CCGCAGATGCAAAATGTCGTACAGAGATTAACCTTGAATTCCTGATGCGCACTCGGAT 900
Db 1177 CCGCAGATGCAAAATGTCGTACAGAGATTAACCTTGAATTCCTGATGCGCACTCGGAT 1236
Qy 901 TTCTGCAAGTCTTTCTGAAAATCTGTGTGAACATCATGCTCTTTTGAACCTTTTGA 960
Db 1237 TTCTGCAAGTCTTTCTGAAAATCTGTGTGAACATCATGCTCTTTTGAACCTTTTGA 1296
Qy 961 GAGCCCAAAACAAAGCCCAAGCCGTCCTTTAGCGGGGGGTCATCTTGGGTTCAAT 1020
Db 1297 GAGCCCAAAACAAAGCCCAAGCCGTCCTTTAGCGGGGGGTCATCTTGGGTTCAAT 1356
Qy 1021 GGTGCGACTCAGAAAGAGTTCTGACTATGTTAAAGAAAGAGACATAAGAGGTGAG 1080
Db 1357 GGTGCGACTCAGAAAGAGTTCTGACTATGTTAAAGAAAGAGACATAAGAGGTGAG 1416
Qy 1081 TTTGAAAGAGACACAGCAAGATTCATTTATCCGAGCCTTGTCAAGCCTTACAGAA 1140
Db 1417 TTTGAAAGAGACACAGCAAGATTCATTTATCCGAGCCTTGTCAAGCCTTACAGAA 1476
Qy 1141 CTGAATTCGGAAGTGTGAGAGCTCTCAGCAGAGACCAAGCCTTACATTTGGAGAGGT 1200
Db 1477 CTGAATTCGGAAGTGTGAGAGCTCTCAGCAGAGACCAAGCCTTACATTTGGAGAGGT 1536
Qy 1201 GCGGAATCTCAGAGGGGCGAGAGCTCGCGAGAGAAAGAAACGAAAGGTTTCCGCGGG 1260
Db 1537 GCGGAATCTCAGAGGGGCGAGAGCTCGCGAGAGAAAGAAACGAAAGGTTTCCGCGGG 1596
Qy 1261 GAGCCGGGGTCCGACCCGAGCCCTGCGCGAGAGAAAGCCCGGGGTAAACAAGCGGCG 1320

Db 1597 GAGCCGGGGTCCGACCCGAGCCCTGCGCGAGAGAAAGCCCGCGGGTTAAACAAGCGGCG 1656
Qy 1321 GAGGAGCGGCTGTGGGCGCCACAGGAGGAAGAGAGAGGTGTTAAGATAGAGCCAG 1380
Db 1657 GAGGAGCGGCTGTGGGCGCCACAGGAGGAAGAGAGAGGTGTTAAGATAGAGCCAG 1716
Qy 1381 CAGAGTAAACCTCAGCCCGCGAGCCAGACAGAGCTCCCTGACTGAGAGTCTTCACTT 1440
Db 1717 CAGAGTAAACCTCAGCCCGCGAGCCAGACAGAGCTCCCTGACTGAGAGTCTTCACTT 1776
Qy 1441 TCCGAGCTGTCTGTGAATCTCGCAGGGGGAGATGAGCCCTGCGCAAGTGAATCTTCCC 1500
Db 1777 TCCGAGCTGTCTGTGAATCTCGCAGGGGGAGATGAGCCCTGCGCAAGTGAATCTTCCC 1836
Qy 1501 AACCTGAGCCCGGACACCAAGAGGCTCTCTCTTGAATCAGCGCGCTGTGAATGACCG 1560
Db 1837 AACCTGAGCCCGGACACCAAGAGGCTCTCTCTTGAATCAGCGCGCTGTGAATGACCG 1896
Qy 1561 GCTTCCCGCGAGCGAGATGAGGATGAGGGCCGAGAGAAAGATTCCAACTGATAAA 1620
Db 1897 GCTTCCCGCGAGCGAGATGAGGATGAGGGCCGAGAGAAAGATTCCAACTGATAAA 1956
Qy 1621 GGTATCTTCAATAGCTTAAGGAAGTGTCTACACAGCGGCAATATCTGAAGATTCGAA 1680
Db 1957 GGTATCTTCAATAGCTTAAGGAAGTGTCTACACAGCGGCAATATCTGAAGATTCGAA 2016
Qy 1681 GTTATCATCTTGAGTTTCAAGACAGTGAAGCAAAAGAGAGCGCATCCGGAAGCACTG 1740
Db 2017 GTTATCATCTTGAGTTTCAAGACAGTGAAGCAAAAGAGAGCGCATCCGGAAGCACTG 2076
Qy 1741 AAAAGTCTATATTCGGAATTTTGAACCTTTGCACAAAATTTCAATATTTTCTCAAG 1800
Db 2077 AAAAGTCTATATTCGGAATTTTGAACCTTTGCACAAAATTTCAATATTTTCTCAAG 2136
Qy 1801 GAAATTTGAGCAACGACTTGGCCCTGTGGGAAGGCGCTCAAAATGCCAATAACAGAGATTAC 1860
Db 2137 GAAATTTGAGCAACGACTTGGCCCTGTGGGAAGGCGCTCAAAATGCCAATAACAGAGATTAC 2196
Qy 1861 CAAGAATTCGCGAGATGATCTGAGAAACATTCAGGCGATGAAGCACTTGGCGGCTCAC 1920
Db 2197 CAAGAATTCGCGAGATGATCTGAGAAACATTCAGGCGATGAAGCACTTGGCGGCTCAC 2256
Qy 1921 CTGTGAAAGCACAGGAGGCTTTGAGGCGCTTGAAGAAATGAATCAAGAGCTCCGCGCG 1980
Db 2257 CTGTGAAAGCACAGGAGGCTTTGAGGCGCTTGAAGAAATGAATCAAGAGCTCCGCGCG 2316
Qy 1981 CTGAGAACTTCTGAGAGACTTTGAGCTGAGAAAGGTGTGTACTTACCGCTCAACACC 2040
Db 2317 CTGAGAACTTCTGAGAGACTTTGAGCTGAGAAAGGTGTGTACTTACCGCTCAACACC 2376
Qy 2041 TTCTCTCTGCGGCACTGACCGGCTCATGCACTTCAAGCAGAGTCTTGAAGCGGCTGTGC 2100
Db 2377 TTCTCTCTGCGGCACTGACCGGCTCATGCACTTCAAGCAGAGTCTTGAAGCGGCTGTGC 2436
Qy 2101 AAAACACACCCGCGAGCCGACGCGACTTCAAGGACTGCGGAGCCGCTTTGGCAGAGATC 2160
Db 2437 AAAACACACCCGCGAGCCGACGCGACTTCAAGGACTGCGGAGCCGCTTTGGCAGAGATC 2496
Qy 2161 ACGAGAGATGAGGCAACAGCTCAAGGTACGATGATCAAGATGAGAGAAATTTCCAGAAAGCTG 2220
Db 2497 ACGAGAGATGAGGCAACAGCTCAAGGTACGATGATCAAGATGAGAGAAATTTCCAGAAAGCTG 2556
Qy 2221 CACGAATCAAGAAAGATTTGATGGAATCTTGTGTGTTCCGGGAAGGAGGTTTCC 2280
Db 2557 CACGAATCAAGAAAGATTTGATGGAATCTTGTGTGTTCCGGGAAGGAGGTTTCC 2616
Qy 2281 ATCTGTCTGGGCAAGCTTCAAGAGCTCTCGGGGAAAGGGCTTCCAGACGCAATGTTCTTC 2340
Db 2617 ATCTGTCTGGGCAAGCTTCAAGAGCTCTCGGGGAAAGGGCTTCCAGACGCAATGTTCTTC 2676
Qy 2341 CTGTTCACACAGTCTCTGCTATACACAGGCGGGGGCTGAAGGCTTCAATCAAGTTTAA 2400
Db 2677 CTGTTCACACAGTCTCTGCTATACACAGGCGGGGGCTGAAGGCTTCAATCAAGTTTAA 2736

OY		2401	GTCACACGGGCAAGCTCCCGCTCTAATGGCATGACATTTGAGAGAGCAGAAACAAGATGGGGG	2460
Db		2737	GTCCACGGGCAAGCTCCCGCTCTAATGGCATGACATTTGAGAGAGCAGAAACAAGATGGGGG	2798
OY		2461	GTGCCCACTGCTGACCTTCCGGGGCCAGCGGCAGTGCATCATCTGTGGCCGCAATTCT	2520
Db		2797	GTGCCCACTGCTGACCTTCCGGGGCCAGCGGCAGTGCATCATCTGTGGCCGCAATTCT	2856
OY		2521	CGGTCCGAGATGAGAGAAGTGGGTTGAGGACATCCAGATGGCCATTGACTGTGGCGGAAG	2580
Db		2857	CGGTCCGAGATGAGAGAAGTGGGTTGAGGACATCCAGATGGCCATTGACTGTGGCGGAAG	2916
OY		2581	AGCAGCACCCCCCGCCCTGAGTTCTGTGGCCAGCACCCCCCTGACAAACAATCCCCTGAT	2640
Db		2917	AGCAGCACCCCCCGCCCTGAGTTCTGTGGCCAGCACCCCCCTGACAAACAATCCCCTGAT	2976
OY		2641	GAAGCCACCGGCGCTGACCAAGAGTCAAGAGATGACTGAGCGCTCGCGACATCCGCT	2700
Db		2977	GAAGCCACCGGCGCTGACCAAGAGTCAAGAGATGACTGAGCGCTCGCGACATCCGCT	3036
OY		2701	GAGCGCCACGGCCCCCGCACCCGGGCAACAACAATGTGCACTGTGTGCGACCCGCAACC	2760
Db		3037	GAGCGCCACGGCCCCCGCACCCGGGCAACAACAATGTGCACTGTGTGCGACCCGCAACC	3096
OY		2761	AGCGCTCCATGTGGGACCTTGACGATCCGACGTGGAGTAATCAGTTGTCTGGAAAACCTGCT	2820
Db		3097	AGCGCTCCATGTGGGACCTTGACGATCCGACGTGGAGTAATCAGTTGTCTGGAAAACCTGCT	3156
OY		2821	AGGAAATTCAAAAACAGCAAGGGGTGGCAGAACTGTGGGTGTGTTCACAACTTCTGC	2880
Db		3157	AGGAAATTCAAAAACAGCAAGGGGTGGCAGAACTGTGGGTGTGTTCACAACTTCTGC	3216
OY		2881	CTGTTCTTCTAACAAATCACACACAGGACATATCCCCCTTGGCACCTGCTCTGGGC	2940
Db		3217	CTGTTCTTCTAACAAATCACACACAGGACATATCCCCCTTGGCACCTGCTCTGGGC	3276
OY		2941	TACTCGCTCACATCCCTCTGAGTCCGAGAAACATCCAGAAAGTACGTCTTCAAAGCTG	3000
Db		3277	TACTCGCTCACATCCCTCTGAGTCCGAGAAACATCCAGAAAGTACGTCTTCAAAGCTG	3336
OY		3001	CAC TTC AAG TCC CCA GGT CT A CT TCA GTG CG GA AG CG A GTA CA C GTT G AA A GG TG G	3060
Db		3337	CAC TTC AAG TCC CCA GGT CT A CT TCA GTG CG GA AG CG A GTA CA C GTT G AA A GG TG G	3396
OY		3061	ATGGAAGTATCCGACGTGCCACCAAGCTCTGCTTGGCAGCCCACTGTGGACCAAAA	3120
Db		3397	ATGGAAGTATCCGACGTGCCACCAAGCTCTGCTTGGCAGCCCACTGTGGACCAAAA	3456
OY		3121	GAGTCTCTTGTGTAT 3135	
Db		3457	GAGTCTCTTGTGTAT 3471	
<hr/>				
RESULT 4				
AAA08582 ID AAA08582 standard; DNA; 4687 BP.				
<hr/>				
XX	AAA08582;			
XX	19-JUL-2000 (first entry)			
DE	Human cytoskeleton associated protein 2 (CYSKP-2) coding sequence.			
XX	KW Cytoskeleton associated protein; CYSKP-2; cancer; proliferative;			
KM	autoimmunity; inflammatory; vesicle trafficking; neurological;			
XX	cardiovascular; cell motility; reproductive; muscle disorder; ss.			
OS	Homo sapiens.			
XX				
Key	Location/Qualifiers			
FH	90..327			
CDL	/+tag= a			
FT				

```

FT      /product= "CYSKP-2"
FN      WO20017355-A2.
PD      30-MAR-2000.
PF      17-SEP-1999;    99MO-US021565.
PR      18-SEP-1998;    98US-0172226P.
PP      27-APR-1999;    99US-0131321P.
PA      (INCY-) INCYTE PHARM INC.
PI      Lal P, Tang YT, Yue H, Hillman JL, Bandman O, Corley NC;
PI      Guegler KJ, Paterson C, Azimzal Y, Baughn MR;
XX      WPI; 2000-283582/24.
DR      P-PsDB; AAY91947.
PT      Human cytoskeleton associated proteins, used to treat cell proliferative,
PT      autoimmune/inflammatory, vesicle trafficking, neurological, cell
PT      motility, reproductive and muscle disorders.
XX      Claim 9; Page 101-102; 113pp; English.
PS      AAA08581-96 encode human cytoskeleton associated proteins 1 to 16 (CYSKP-
CC      1 to CYSKP-16) respectively. The sequences can be used to treat and
CC      diagnose cancer and cell proliferative, autoimmune/inflammatory, vesicle
CC      trafficking, neurological, cardiovascular, cell motility, reproductive
CC      and muscle disorders. Pharmaceutical compositions containing CYSKP-2 to
CC      CYSKP-16 can be used to treat or prevent disorders associated with
CC      decreased expression or activity of CYSKP (claimed), for example,
CC      atherosclerosis, cirrhosis, hepatitis, myelofibrosis, psoriasis, cancers,
CC      autoimmune/antidifferentiation disorders such as allergies, anemia, asthma,
CC      acquired immunodeficiency syndrome (AIDS), Crohn's disease, diabetes,
CC      mellitus, Goodpasture's syndrome, osteoarthritis, scleroderma and trauma.
CC      CYSKP antagonists can be used to treat or prevent a disorder associated
CC      with increased expression or activity of CYSKP (claimed).
SQ      Sequence 4687 BP; 1179 A; 1261 C; 1196 G; 1051 T; 0 U; 0 Other;

Query Match          99.9%; Score 3131.8; DB 3; Length 4687;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3133; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

        1 ATGGGAGAAAATAGAGCAGAGGCCCGACCATCAAGACTGTGGGGCCCCGGAAAAATTTCG   60
Dd      90 ATGGGAGAAATATAGAGCAGAGGCCCGACCATCAAGACTGTGGGGCCCCGGAAAAATTTCG   149
Oy      61 GGAGTACAGTACTCTTGGAACGTGACAGAAAGCCGCCCCCAACACTTCAGAAAACTCGTG   120
Dd      150 GGAGTACAGTACTCTTGGAACGTGACAGAAAGCCGCCCCCAACACTTCAGAAAACTCGTG   209
Oy      121 TTCATCAAAAATCCAGATGCTGTGAGAACCCAGAGAGGATTGGAAGTTCCCAAAGACT       180
Dd      210 TCATCAAAAATCCAGATGCTGTGAGAACCCAGAGAGGATTGGAAGTTCCCAAAGACT       269
Oy      181 CCTGGGAAAGGTGCTGCTGATGCAGATTGTAACAACCACTCAACTGTGGAAGGTGACTAT     240
Dd      270 CCTGGGAAAGGTGCTGCTGATGCAGATTGTAACAACCACTCAACTGTGGAAGGTGACTAT     329
Oy      241 TTGGGCTCCAGATTCTCTGATCAAAAAAAGTACCGGTGTGGCTGATCTCTTAAACC       300
Dd      330 TTGGGCTCCAGATTCTCTGATCAAAAAAAGTACCGGTGTGGCTGATCTCTTAAACC       389
Oy      301 ATTGTGAAAACGATTTAGAAGGCCCAAAGCACTGTGTGTTAAGTTTGGTGTGAAATTTCTT     360
Dd      390 ATTGTGAAAACGATTTAGAAGGCCCAAAGCACGTTGTGTGAAGTTTGTGTGAATTTCTT     449
Oy      361 CGGCTTAGACACACAACTTCCAAGAAAGAACTCAAAAGTACTGTTCGCTGCAGAGTG       420
Dd      450 CGGCTTAGACACACAACTTCCAAGAAAGAACTCAAAAGTACTGTTCGCTGCAGAGTG       509

```

421 AAGCAGACCTGGCTCAAGCAGGTTGAGCTGTAATGACACACGACGACCTCTTGATT 480
| | | | |
Db 510 AAGCAGACCTGGCTCAAGCAGGTTGAGCTGTAATGACACACGACGACCTCTTGATT 569
| | | | |
Qy 481 TCACACATTTGCAATCTGAGATTTGGGATTTTGTATGAAGCTTGGACAGAGACCTTA 540
| | | | |
Db 570 TCACACATTTGCAATCTGAGATTTGGGATTTTGTATGAAGCTTGGACAGAGACCTTA 629
| | | | |
Qy 541 GGAATAAATAATATCTACCTGACGAGACGCACTGAGAGACAAATTCGTGAATTTTCA 600
| | | | |
Db 630 GGAATAAATAATATCTACCTGACGAGACGCACTGAGAGACAAATTCGTGAATTTTCA 689
| | | | |
Qy 601 CATTAACCATTTGACAAACACGACGAGATCAGATTTCCAGCTCTTAGAGATGGCCGT 660
| | | | |
Db 690 CATTAACCATTTGACAAACACGACGAGATCAGATTTCCAGCTCTTAGAGATGGCCGT 749
| | | | |
Qy 661 CGGCTAGAGATGTATGAAATCCGGTTGCAACCGGCCAAGACAGAGAGCAGAAAGATC 720
| | | | |
Db 750 CGGCTAGAGATGTATGAAATCCGGTTGCAACCGGCCAAGACAGAGAGCAGAAAGATC 809
| | | | |
Qy 721 AATCTGGCCGTTGGCCAAACGCGGAATTTAGTGTTCAGGGTTTCACTAAGATCAATGCG 780
| | | | |
Db 810 AATCTGGCCGTTGGCCAAACGCGGAATTTAGTGTTCAGGGTTTCACTAAGATCAATGCG 869
| | | | |
Qy 781 TTCAACTGGGCCAAGGTGCGAAGCTGAGCTTCAAGAGAGAGCCTTCTCAATCAAGCTC 840
| | | | |
Db 870 TTCAACTGGGCCAAGGTGCGAAGCTGAGCTTCAAGAGAGAGCCTTCTCAATCAAGCTC 929
| | | | |
Qy 841 CGGCCAGATGCCAATATGTCGTACAGAGATACCTTGAATTCCTGATGCGCAGTGGGAT 900
| | | | |
Db 930 CGGCCAGATGCCAATATGTCGTACAGAGATACCTTGAATTCCTGATGCGCAGTGGGAT 989
| | | | |
Qy 901 TTCTGAAATCTCTTGGGAAAATCTGTGTGAACATCAGCTCTTGTGAATTTTGA 960
| | | | |
Db 990 TTCTGAAATCTCTTGGGAAAATCTGTGTGAACATCAGCTCTTGTGAATTTTGA 1049
| | | | |
Qy 961 GAGCCCAACCAAGCCGCTCTTTAGCCGGGGGTCAATTCGTTGAGT 1020
| | | | |
Db 1050 GAGCCCAACCAAGCCGCTCTTTAGCCGGGGGTCAATTCGTTGAGT 1109
| | | | |
Qy 1021 GGTGGACTGAGAGAGGTTCTGACTATGTTAAAGAGAGACATTAAGAGGTGACG 1080
| | | | |
Db 1110 GGTGGACTGAGAGAGGTTCTGACTATGTTAAAGAGAGACATTAAGAGGTGACG 1169
| | | | |
Qy 1081 TTTGAAAGAGAGACAGAAAGATTCAATTCATCCGAGCCTTGCTTCAAGCTTACAGAA 1140
| | | | |
Db 1170 TTTGAAAGAGAGACAGAAAGATTCAATTCATCCGAGCCTTGCTTCAAGCTTACAGAA 1229
| | | | |
Qy 1141 CTGAATTCGAAAGTGTGAGAGCTCTCAGCAGAGACACGACCTTACATTTGAGAGAGT 1200
| | | | |
Db 1230 CTGAATTCGAAAGTGTGAGAGCTCTCAGCAGAGACACGACCTTACATTTGAGAGAGT 1289
| | | | |
Qy 1201 GCCGAAATCTCAGGGGGCCAGAGCTCCGCGAGAGAAAGAACCGAAGTTTCCCGGG 1260
| | | | |
Db 1290 GCCGAAATCTCAGGGGGCCAGAGCTCCGCGAGAGAAAGAACCGAAGTTTCCCGGG 1349
| | | | |
Qy 1261 GAGCCGGGGTGCACCCGAGCCCTGCGCGAGAGAGAAAGCCCGCGGGTAAACAAGAGGG 1320
| | | | |
Db 1350 GAGCCGGGGTGCACCCGAGCCCTGCGCGAGAGAGAAAGCCCGCGGGTAAACAAGAGGG 1409
| | | | |
Qy 1321 GACGAGAGCCGCTCGCGGCCACGAGAGAGAGAGAGAGTGTGTAAGATAGAGCCAG 1380
| | | | |
Db 1410 GACGAGAGCCGCTCGCGGCCACGAGAGAGAGAGAGAGTGTGTAAGATAGAGCCAG 1469
| | | | |
Qy 1381 CAGAGTAAACTTGAAGCCCCGACAGCAGCAGGCTCCCTGACTGGCAGTCTTCACTTT 1440
| | | | |
Db 1470 CAGAGTAAACTTGAAGCCCCGACAGCAGCAGGCTCCCTGACTGGCAGTCTTCACTTT 1529
| | | | |
Qy 1441 TCCGAGCTGTCTGTAACTGCAAGGGGGAGTGGCCCTGCGCAAGTGAATTTGTCTCC 1500
| | | | |
Db 1530 TCCGAGCTGTCTGTAACTGCAAGGGGGAGTGGCCCTGCGCAAGTGAATTTGTCTCC 1589
| | | | |
Qy 1501 AACCTGAGCCCGCAGACAAAGCAGGCTCTCCCTTGATCAGCCCGCTGTGATGACAG 1560
| | | | |

1590 AACCTGAGCCCGCAGACCAAGAGGCTCTCCCTTGATCAGCCGCTGTGATGACAG 1649
| | | | |
Qy 1561 GCTTCCCCCGGACGGAAGATGAGAGTGAAGGGCCGGAAGAAAGATTTCCCACTGATAA 1620
| | | | |
Db 1650 GCTTCCCCCGGACGGAAGATGAGAGTGAAGGGCCGGAAGAAAGATTTCCCACTGATAA 1709
| | | | |
Qy 1621 GGTACTTCAATGACTAAGAAAGTGTCTACACCGAGCGCAACATATCTGAAGATCTGNA 1680
| | | | |
Db 1710 GGTACTTCAATGACTAAGAAAGTGTCTACACCGAGCGCAACATATCTGAAGATCTGNA 1769
| | | | |
Qy 1681 GTTATCACTTGTGTGTTTCAAGACAGTGAAGCAAGAGAGAGCGCATCCGAGACCTG 1740
| | | | |
Db 1770 GTTATCACTTGTGTGTTTCAAGACAGTGAAGCAAGAGAGAGCGCATCCGAGACCTG 1829
| | | | |
Qy 1741 AAAAGTCTCATATTTCCGGAATTTTGAACCTTTGCAAAATTTCAATCAATTTTCTCAAG 1800
| | | | |
Db 1830 AAAAGTCTCATATTTCCGGAATTTTGAACCTTTGCAAAATTTCAATCAATTTTCTCAAG 1889
| | | | |
Qy 1801 GAAATTTGAGCAACGACTTGGCCCTGTGGAGAGGCGCCTCAAAATGCCCAATCAGAGATTAC 1860
| | | | |
Db 1890 GAAATTTGAGCAACGACTTGGCCCTGTGGAGAGGCGCCTCAAAATGCCCAATCAGAGATTAC 1949
| | | | |
Qy 1861 CAAGAATTCGGGATGTATGCTGAAGAAATTCAGGGCATGAAAGCACCTGGCGGCTCAC 1920
| | | | |
Db 1950 CAAGAATTCGGGATGTATGCTGAAGAAATTCAGGGCATGAAAGCACCTGGCGGCTCAC 2009
| | | | |
Qy 1921 CTGTGGAAGCACAGAGAGGCTTGAAGGCTCTGAGAAATGGAATGAAGATCTCCGGGG 1980
| | | | |
Db 2010 CTGTGGAAGCACAGAGAGGCTTGAAGGCTCTGAGAAATGGAATGAAGATCTCCGGGG 2069
| | | | |
Qy 1981 CTGAGAACTTCTGAGAGACTTTGAGCTGCAAGAGTGTGTAACTTCAACCGCTCAACAC 2040
| | | | |
Db 2070 CTGAGAACTTCTGAGAGACTTTGAGCTGCAAGAGTGTGTAACTTCAACCGCTCAACAC 2129
| | | | |
Qy 2041 TTCTCTTGGGCACTGCAACCGGCTCATGCACTAACAGAGTCTTGAAGCGCTGTGTC 2100
| | | | |
Db 2130 TTCTCTTGGGCACTGCAACCGGCTCATGCACTAACAGAGTCTTGAAGCGCTGTGTC 2189
| | | | |
Qy 2101 AAACACACCCGCGCAGACGACGCGCACTTTCAGGGGCTCCGAGCCGCTTTGGCAGAGATC 2160
| | | | |
Db 2190 AAACACACCCGCGCAGACGACGCGCACTTTCAGGGGCTCCGAGCCGCTTTGGCAGAGATC 2249
| | | | |
Qy 2161 ACGAGATGATGAGCAGACTCCACGATGATGATCAAGATGAGAAATTTCCAGAAAGCTG 2220
| | | | |
Db 2250 ACGAGATGATGAGCAGACTCCACGATGATGATCAAGATGAGAAATTTCCAGAAAGCTG 2309
| | | | |
Qy 2221 CACGAATCTGAAAGATTTGATTTGCAATTTGTCATTTGTTTCCGGAAGGAGATTTC 2280
| | | | |
Db 2310 CACGAATCTGAAAGATTTGATTTGCAATTTGTCATTTGTTTCCGGAAGGAGATTTC 2369
| | | | |
Qy 2281 ATCCGTCGAGGAGCTGACAGAGCTCTCGGGGAAAGGGGCTCCAGACGACATGTTCTTC 2340
| | | | |
Db 2370 ATCCGTCGAGGAGCTGACAGAGCTCTCGGGGAAAGGGGCTCCAGACGACATGTTCTTC 2429
| | | | |
Qy 2341 CTGTTCAAAGAGCTCTGCTATACAGAGCGGGGGCTGACGCGCTCCAAATCAGTTTAA 2400
| | | | |
Db 2430 CTGTTCAAAGAGCTCTGCTATACAGAGCGGGGGCTGACGCGCTCCAAATCAGTTTAA 2489
| | | | |
Qy 2401 GTTCACGGGCAAGCTCCGCTCTATGAGCATGACGATTTGAGAGAGCGAAAGCAGATGGGG 2460
| | | | |
Db 2490 GTTCACGGGCAAGCTCCGCTCTATGAGCATGACGATTTGAGAGAGCGAAAGCAGATGGGG 2549
| | | | |
Qy 2461 GTGGCCCACTGCTGACCTCCGCGGCAAGGGGCAAGTGCATCATGCTGGCGCCAGTTCT 2520
| | | | |
Db 2550 GTGGCCCACTGCTGACCTCCGCGGCAAGGGGCAAGTGCATCATGCTGGCGCCAGTTCT 2609
| | | | |
Qy 2521 CGGTCCAGATGAGAAATGGTGTGAGACATTCAGATGCGCAATTTGACTGGCGGAGAG 2580
| | | | |
Db 2610 CGGTCCAGATGAGAAATGGTGTGAGACATTCAGATGCGCAATTTGACTGGCGGAGAG 2669
| | | | |
Qy 2581 AGCAGAGCCCGCCCTGAGTTCTTGGCCAGAGCCCGCTGACACAGATCCCTTGAAT 2640
| | | | |

Db 2670 AGCAGACGCCCCCTGAGTTCCTG6CCAGACGCCCCCTGCAACAAGTCCCTGAT 2729
Qy 2641 GAAGCCACCGCGCTGACCAAGATGACCTGAGCCCTCGGCCAATGCTG 2700
Db 2730 GAAGCCACCGCGCTGACCAAGATGACCTGAGCCCTCGGCCAATGCTG 2789
Qy 2701 GAGCGCAGGCGCGCGCAGCAGCAATGATGCTGCTGCGCAGCCCAACAC 2760
Db 2790 GAGCGCAGGCGCGCGCAGCAGCAATGATGCTGCTGCGCAGCCCAACAC 2849
Qy 2761 AGCGTCTCCATGATGATCTTACGATGCAAGTGAAGATGATGCTGCTG 2820
Db 2850 AGCGTCTCCATGATGATCTTACGATGCAAGTGAAGATGATGCTGCTG 2909
Qy 2821 AGGAATTTCAAAACAGCAACGCGTGGCAGAACCTGCGGTTCACAACTTCTGC 2880
Db 2910 AGGAATTTCAAAACAGCAACGCGTGGCAGAACCTGCGGTTCACAACTTCTGC 2969
Qy 2881 CTGTTCTTCAAAATCAACAGCAATCATCCCTTSCAGCCTGCTGCTGCGC 2940
Db 2970 CTGTTCTTCAAAATCAACAGCAATCATCCCTTSCAGCCTGCTGCTGCGC 3029
Qy 2941 TACTGCTCAACATCCCTCTGATGCTCGAGAACATCCAGAAAGATGATG 3000
Db 3030 TACTGCTCAACATCCCTCTGATGCTCGAGAACATCCAGAAAGATGATG 3089
Qy 3001 CACTTCAAGTCCCAAGTCTTACTTCAAGGCGGAAAGCAATGATGATG 3060
Db 3090 CACTTCAAGTCCCAAGTCTTACTTCAAGGCGGAAAGCAATGATGATG 3149
Qy 3061 ATGGAAGTATCCGAGTGCACAGCTGCTGCGGCAACCCAGCTGTTGAGCA 3120
Db 3150 ATGGAAGTATCCGAGTGCACAGCTGCTGCGGCAACCCAGCTGTTGAGCA 3209
Qy 3121 GAGTCTCTTGTGTAT 3135
Db 3210 GAGTCTCTTGTGTAT 3224

RESULT 5
AAS64760
ID AAS64760 standard; cDNA; 3187 BP.
XX AAS64760;
AC AAS64760;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #564.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR P-PsDB; ABG00573.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnosis, forensic, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess

PT biodiversity.
XX
XX Claim 1; SEQ ID NO 564; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridization probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 3187 BP; 810 A; 874 C; 876 G; 627 T; 0 U; 0 Other;
Query Match 99.5%; Score 3119.2; DB 5; Length 3187;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3132; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
Qy 1 ATGGGAAATATAGAGAGAGCGCCAGCCAGATGACAGTGGGGCCCCGGAATTTGG 60
Db 49 ATGGGAAATATAGAGAGAGCGCCAGCCAGATGACAGTGGGGCCCCGGAATTTGG 108
Qy 61 GGGATCAGTACCTTGGAAAGTGGACAGAGCGCCCCCAACACTTCAGAAATTCGTG 120
Db 109 GGGATCAGTACCTTGGAAAGTGGACAGAGCGCCCCCAACACTTCAGAAATTCGTG 168
Qy 121 TCCATCAAAATCCAGATGCTGATGACACCCAGAGAGCATTTGAACTTCACAAAGAGCT 180
Db 169 TCCATCAAAATCCAGATGCTGATGACACCCAGAGAGCATTTGAACTTCACAAAGAGCT 228
Qy 181 CCTGGGAAGGTGCTGCTGATGATGCAAGTTGCAACCACTCACTCTGGAAGTGA 240
Db 229 CCTGGGAAGGTGCTGCTGATGATGCAAGTTGCAACCACTCACTCTGGAAGTGA 288
Qy 241 TTTGGCTCGAGTTTCTGATCAAAAAGATCAGCGGTGCTGATCTCTTAAACC 300
Db 289 TTTGGCTCGAGTTTCTGATCAAAAAGATCAGCGGTGCTGATCTCTTAAACC 348
Qy 301 ATTTGAAAACAGATTAGAGGCCAAAGCAAGTTTGTAGTTTGTGTAATTCCTT 360
Db 349 ATTTGAAAACAGATTAGAGGCCAAAGCAAGTTTGTAGTTTGTGTAATTCCTT 408
Qy 361 CGGCTGACCAACACAACTCCAAAGAAATCAACAAGTACTGTTCCGCTGCAAGTG 420
Db 409 CGGCTGACCAACACAACTCCAAAGAAATCAACAAGTACTGTTCCGCTGCAAGTG 468
Qy 421 AAGCAGAGCTTGGCTCAAGGAGTTGACGTGTAATGACACGAGCAGACTCTCTGATT 480
Db 469 AAGCAGAGCTTGGCTCAAGGAGTTGACGTGTAATGACACGAGCAGACTCTCTGATT 528
Qy 481 TCACACATTTGCAATCTGAGATTGGGATTTTGAATGAACTTGGACAGAGACCTTA 540
Db 529 TCACACATTTGCAATCTGAGATTGGGATTTTGAATGAACTTGGACAGAGACCTTA 588
Qy 541 GCAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 600
Db 589 GCAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 648
Qy 601 CATTAACCATTTGACAAACACAGCAGATGATGATTTCCAGCTCTAGAGATTGCCGT 660

Db 649 CATPACCAATTGGAACAAACACAGAGATCATGATTTCCAGTCTCTAAGATTGCCGT 708
Qy 661 CGGCTAGAGATGTTATGGAATCCGGTTGCACCCGGCCAAAGACAGGGAAGCCAGGAAGATC 720
Db 709 CGGCTAGAGATGTTATGGAATCCGGTTGCACCCGGCCAAAGACAGGGAAGCCAGGAAGATC 768
Qy 721 AATTCGGCCGTTGCGAACAAGGGAATTCATAGTGTTCAGGGTTTCACTAAGATCAATGCC 780
Db 769 AATCTGGCCGTTGCGAACAAGGGAATTCATAGTGTTCAGGGTTTCACTAAGATCAATGCC 828
Qy 781 TTCAACTGGGCCCAAGGTGCGGAAGCTGAGCTTCAAGAGAAGCGCTTTCTCATCAAGCTC 840
Db 829 TTCAACTGGGCCCAAGGTGCGGAAGCTGAGCTTCAAGAGAAGCGCTTTCTCATCAAGCTC 888
Qy 841 CGGCCAGATGCCAATATGTCGTAACAGATATCTTGAATTCCTGATGCGCAGTCGGGAT 900
Db 889 CGGCCAGATGCCAATATGTCGTAACAGATATCTTGAATTCCTGATGCGCAGTCGGGAT 948
Qy 901 TTCTGCAAGTCTTCTGGAATAATCTGTGTGAACATCATGCTTCTTGAATCTTTTGA 960
Db 949 TTCTGCAAGTCTTCTGGAATAATCTGTGTGAACATCATGCTTCTTGAATCTTTTGA 1008
Qy 961 GAGCCCAACCAAAAGCCCAAGCCGTCCTTTAGCCGGGGTCATCATTTGGTTCACT 1020
Db 1009 GAGCCCAACCAAAAGCCCAAGCCGTCCTTTAGCCGGGGTCATCATTTGGTTCACT 1068
Qy 1021 GGTCCGATCTCAAGAGAGGTTCTCGACTATGTTAAAGAGAGACATAAGAGGTGCAG 1080
Db 1069 GGTCCGATCTCAAGAGAGGTTCTCGACTATGTTAAAGAGAGACATAAGAGGTGCAG 1128
Qy 1081 TTTGAAAGAGACAGCAAGATTCATTTATCCGAGCCTTGCTTCAAGCCTACAGAA 1140
Db 1129 TTTGAAAGAGACAGCAAGATTCATTTATCCGAGCCTTGCTTCAAGCCTACAGAA 1188
Qy 1141 CTGAATTCGAAAGTGTGAGAGAGTCTAGCAGAGACCAAGCCTTACATTTGAGAGA -AGG 1199
Db 1189 CTGAATTCGAAAGTGTGAGAGAGTCTAGCAGAGACCAAGCCTTACATTTGAGAGAG 1248
Qy 1200 TGCCTGAATCTCCAGGGGGCCAGAGCTGCCGGCGAGAAAGAACCCGAAGTTTCCGCCG 1259
Db 1249 TGCCTGAATCTCCAGGGGGCCAGAGCTGCCGGCGAGAAAGAACCCGAAGTTTCCGCCG 1308
Qy 1260 GGAAGCCGGGGTGCACCCGAGCCCTGCGCGAGAGAACCCCGGGGTAAACAAGAGGC 1319
Db 1309 GGAAGCCGGGGTGCACCCGAGCCCTGCGCGAGAGAACCCCGGGGTAAACAAGAGGC 1368
Qy 1320 GGAAGAGCCGCTCGGCGCCCAAGAGAAAGAGAGAGTGTGTTAAGATAGACCCA 1379
Db 1369 GGAAGAGCCGCTCGGCGCCCAAGAGAAAGAGAGAGTGTGTTAAGATAGACCCA 1428
Qy 1380 GCAAGATTAACCTCAGCCCCCGAGCCCAAGACAGAGCTCCCTGATCTGAGATCTCACT 1439
Db 1429 GCAAGATTAACCTCAGCCCCCGAGCCCAAGACAGAGCTCCCTGATCTGAGATCTCACT 1488
Qy 1440 TTCGAGAGTGTGTGTAATCGCAGGGGGAGTGGCCCTGCAAGCTGATCCTTGTCTCC 1499
Db 1489 TTCGAGAGTGTGTGTAATCGCAGGGGGAGTGGCCCTGCAAGCTGATCCTTGTCTCC 1548
Qy 1500 CAACCTGAGCCCCGACCAAGAGAGGCTCTCCCTTGAATCAAGCCGCTGTAATGACCA 1559
Db 1549 CAACCTGAGCCCCGACCAAGAGAGGCTCTCCCTTGAATCAAGCCGCTGTAATGACCA 1608
Qy 1560 GGCCTTCGCCCCCGAGCGAATGAAGATGAAGGCGCGAAGAAAGATTTCCCACTGATTA 1619
Db 1609 GGCCTTCGCCCCCGAGCGAATGAAGATGAAGGCGCGAAGAAAGATTTCCCACTGATTA 1668
Qy 1620 AGGCTACTCATAGTAAGTAAGTGTCTACCAAGCGCAACATATCTGAAGGATCTCGA 1679
Db 1669 AGGCTACTCATAGTAAGTAAGTGTCTACCAAGCGCAACATATCTGAAGGATCTCGA 1728
Qy 1680 AGTTATCACTTGTGTTTCAAGACAAGTGAACAAAGAGAGCCATGCCGGAAGCACT 1739

Db 1729 AGTTATCACTTGTGTTTCAAGACAAGTGAACAAAGAGAGCCCATGCCGAAGCACT 1788
Qy 1740 GAAAAGTCTCATATCCCGAATTTTGAACCTTGAACAAATTTCTAATATTTCTCA 1799
Db 1789 GAAAAGTCTCATATCCCGAATTTTGAACCTTGAACAAATTTCTAATATTTCTCA 1848
Qy 1800 GGAATTTGAGCAACGACTTGGCCCTGAGGAAGCCGCTCAATATGCCAAATCAGAGATTA 1859
Db 1849 GGAATTTGAGCAACGACTTGGCCCTGAGGAAGCCGCTCAATATGCCAAATCAGAGATTA 1908
Qy 1860 CCAAGAAATCGCGATGTCTGTAAGAAACATTCAGGCAATGAAGCACTGGCCGCTCA 1919
Db 1909 CCAAGAAATCGCGATGTCTGTAAGAAACATTCAGGCAATGAAGCACTGGCCGCTCA 1968
Qy 1920 CCTGTGGAAGCAACAGAGGCTTGAAGCCCTGGAAGATGAAGATCAAGAGCTCCGGCG 1979
Db 1969 CCTGTGGAAGCAACAGAGGCTTGAAGCCCTGGAAGATGAAGATCAAGAGCTCCGGCG 2028
Qy 1980 GCTGGAAGAACTTTCAGAGACTTTCAGCTGAGAGGTGTGTACTTACCGCTCAACAC 2039
Db 2029 GCTGGAAGAACTTTCAGAGACTTTCAGCTGAGAGGTGTGTACTTACCGCTCAACAC 2088
Qy 2040 CTTCCTCTGCGGCCATCTGCACCCGCTCATGCACTAACAGAGTCTGAGCCGCTGTG 2099
Db 2089 CTTCCTCTGCGGCCATCTGCACCCGCTCATGCACTAACAGAGTCTGAGCCGCTGTG 2148
Qy 2100 CAAACACCAACCGCGAGCAACGCGCATTCAGGGAGTCCGAGCCGCTTGGCAGAGAT 2159
Db 2149 CAAACACCAACCGCGAGCAACGCGCATTCAGGGAGTCCGAGCCGCTTGGCAGAGAT 2208
Qy 2160 CACGAGATGTGTGACAGCTTCAAGTACATGATCAAGATGGAATAATTTCCAGAGCT 2219
Db 2209 CACGAGATGTGTGACAGCTTCAAGTACATGATCAAGATGGAATAATTTCCAGAGCT 2268
Qy 2220 GCAAGAACTCAAGAAATTTGATTTGCAATCATCTTGTGTTCCGGGAAGGAGATT 2279
Db 2269 GCAAGAACTCAAGAAATTTGATTTGCAATCATCTTGTGTTCCGGGAAGGAGATT 2328
Qy 2280 CATCCGTGCGGCAACCTCAGCAAGCTCCGGGGAAAGGGGCTCCAGAGCCGATGTTCTT 2339
Db 2329 CATCCGTGCGGCAACCTCAGCAAGCTCCGGGGAAAGGGGCTCCAGAGCCGATGTTCTT 2388
Qy 2340 CCTGTTCAACGAGCTCTGCTATACACAGAGCCGGGGTGAACGCGCTCCAACTCACTTAA 2399
Db 2389 CCTGTTCAACGAGCTCTGCTATACACAGAGCCGGGGTGAACGCGCTCCAACTCACTTAA 2448
Qy 2400 AGTCCAGCGGCACTCCGCTCTATGCAATGAAGATGAAGAGACCAAGAGTGGGG 2459
Db 2449 AGTCCAGCGGCACTCCGCTCTATGCAATGAAGATGAAGAGACCAAGAGTGGGG 2508
Qy 2460 GGTGCGCCCACTGCGTGAACCTCCGGGGCAAGGCGAGTCAATCATGTCGCGCCAGTTT 2519
Db 2509 GGTGCGCCCACTGCGTGAACCTCCGGGGCAAGGCGAGTCAATCATGTCGCGCCAGTTT 2568
Qy 2520 TCGGTCCGAGATGGAAGTGGGTGAAGACATCCAGATGGCCATTGACCTTGGCGAGAA 2579
Db 2569 TCGGTCCGAGATGGAAGTGGGTGAAGACATCCAGATGGCCATTGACCTTGGCGAGAA 2628
Qy 2580 GAGCAGCAGCCCCCGCTGAGTTCCTGGCCAGAGCCCCCTTGACCAACAGTCCCCCTGA 2639
Db 2629 GAGCAGCAGCCCCCGCTGAGTTCCTGGCCAGAGCCCCCTTGACCAACAGTCCCCCTGA 2688
Qy 2640 TGAAGCACAAGCGGTGACAGAGATCAGAGATCACTGAGCGCTCGCGCACTTCGCT 2699
Db 2689 TGAAGCACAAGCGGTGACAGAGATCAGAGATCACTGAGCGCTCGCGCACTTCGCT 2748
Qy 2700 GGAAGCGCAAGGCCCGGCAACCGGCGCAACAAATGATGACGTCGCGCAACCGCAAC 2759
Db 2749 GGAAGCGCAAGGCCCGGCAACCGGCGCAACAAATGATGACGTCGCGCAACCGCAAC 2808
Qy 2760 CAGGCTCTCATGATGGAATTCAGCATGCAAGTGAAGATCAGTTGTCTGGAACCTGCT 2819
Db 2809 CAGGCTCTCATGATGGAATTCAGCATGCAAGTGAAGATCAGTTGTCTGGAACCTGCT 2868

Qy		2820	GAGGAAATTCAAAAACACAACGGGTGCAGAAGCTGGTGGTGTTCACAAATTCTCG	2879
Dd		2869	GAGGAATAATTCAAAACACGAAACGGGTGGCAGAAGCTGGTGGTGTTCACAAATTCTCG	2928
Qy		2880	CCTGTTCTTTCACAAATCACACAGGAAMATCATCCCTTGCACGCTGCTGCTCGG	2933
Dd		2929	CCTGTTCTTTCACAAATCACACAGGAAMATCATCCCTTGCACGCTGCTGCTCGG	2988
Qy		2940	CTACTCGCTCACCATCCCTCTGAGTCCGAGAACATCCAGAAAGCTACGTGTTCAAGCT	2999
Dd		2989	CTACTCGCTCACCATCCCTCTGAGTCCGAGAACATCCAGAAAGCTACGTGTTCAAGCT	3048
Qy		3000	GCACCTCAAGCCCAACGCTCTACTCTTAGGGCCGAAAGCCAGATCAGCTTCGAAAGGTG	3055
Dd		3049	GCACCTCAAGTCCCAACGCTCTACTCTTAGGGCCGAAAGCCAGATCAGCTTCGAAAGGTG	3108
Qy		3060	GATGGAAGTATCCGCAAGTGCACCAGCTCTGCTCCGAGACCACCACTGTTGAGCCACA	3119
Dd		3109	GATGGAAGTATCCGCAAGTGCACCAGCTCTGCTCCGAGACCACCACTGTTGAGCCACA	3168
Qy		3120	AGAGTCTCTTGTGTAT	3135
Dd		3169	AGAGTCTCTTGTGTAT	3184

RESULT 6
AAC98992
ID AAC98992 standard; cDNA; 3094 BP.
XX
AC AAC98992;
XX
DT 09-MAR-2001 (first entry)
DE Human pancreatic cancer antigen nucleotide sequence SEQ ID NO:220.
XX
KW Human; pancreas; pancreatic cancer; pancreatic cancer antigen; detection;
KM diagnosis; identification; cyrostatic; neuroprotective; nootropic;
KW immunomodulatory; relaxant; contraceptive; gynaecological;
KW antitumefactor; cardiant; gene therapy; chromosome mapping;
KW linkage analysis; tissue identification; class typing; forensic; neural;
KM immune system; muscular; reproductive; gastrointestinal; pulmonary;
KM cardiovascular; renal; proliferative; ss.
XX
OS Homo sapiens.
XX
PN WO200055320-A1.
PD 21-SEP-2000.
PF 08-MAR-2000; 2000WO-US005989.
PR 12-MAR-1999; 99US-0124270P.
PA (HUMA-) HUMAN GENOME SCI INC.
PS
PI Rosen CA, Ruben SM;
DR WPI; 2000-579444/54.
XX P-PADB; AAB54227.

New nucleic acid that is a pancreatic cancer antigen for preventing,
treating, or ameliorating a medical condition, particularly pancreatic
cancer, or for use in assays for diagnosing a pathological condition.

claim 1; Page 664-665; 1379pp; English.

AAC98773 to AAC99231 encode the human pancreatic cancer associated
proteins, called pancreatic cancer antigens, given in AAB54008 to
AAB54466. The human pancreatic cancer antigens have cytostatic,
neuroprotective, nootropic, immunomodulatory, relaxant, contraceptive,
gynaecological, cardiant and antiinflammatory activities, and can be used
in gene therapy. The polynucleotide and proteins can be used for

Query Match	45.7%	Score 1431.2	DB 3	Length 3094
Best Local Similarity	99.8%	Pred. No. 0		
Matches 1441	Conservative	2	Mismatches	0
			Indels	1
			Gaps	1
Qy	1692	GTGGTTTCAGAGCA	CAGTGAAGCAAAAGAGAGCCATGCCGGAACCTGAAGATTCAT	1755
Db	62	GTGGTTTCAGAGCA	CAGTGAAGCAAAAGAGAGCCATGCCGGAACCTGAAGATTCAT	121
Qy	1752	ATTCGCGAATTTTAA	CCCTTTGCAAAATTTCTACTAATTTTTCAGAGAAATTTGAGCA	1811
Db	122	ATTCGCGAATTTTAA	CCCTTTGCAAAATTTCTACTAATTTTTCAGAGAAATTTGAGCA	181
Qy	1812	ACGACTTGCCCTGTG	GGAAGGCGCCTCAATGCCCAATCAGAGATTACCAAGAAATCGG	1871
Db	182	ACGACTTGCCCTGTG	GGAAGGCGCCTCAATGCCCAATCAGAGATTACCAAGAAATCGG	241
Qy	1872	CGATGTCAATGCTGA	AAACATTTCAAGGCAATGACACTGGCGGCTCACTGTGGAAGCA	1933
Db	242	CGATGTCAATGCTGA	AAACATTTCAAGGCAATGACACTGGCGGCTCACTGTGGAAGCA	301
Qy	1932	CAGGAGAGCCCTGG	AGGCGCCTGGAGAAATGGAATCAAGAGCTCCGGGGGGCTGGAGAACTT	1991
Db	302	CAGGAGAGCCCTGG	AGGCGCCTGGAGAAATGGAATCAAGAGCTCCGGGGGGCTGGAGAACTT	361
Qy	1992	CTGCAAGAGACTTTG	AGCTGCAAGAGAGTGTCTTCACTACCGCTCAACACTTCTCTCTCGG	2051
Db	362	CTGCAAGAGACTTTG	AGCTGCAAGAGAGTGTCTTCACTACCGCTCAACACTTCTCTCTCGG	421
Qy	2052	GCCACTGACCCGG	CTCAATGCACTACAGACAGGCTCTGAGAGCGCTGTGCAAAACCAACC	2111
Db	422	GCCACTGACCCGG	CTCAATGCACTACAGACAGGCTCTGAGAGCGCTGTGCAAAACCAACC	481
Qy	2112	GCCGAGCAGCGCG	CACTTCAAGGACCTCCGACCGCCTTTGGCAGAGATCAACGGAGATGCT	2171
Db	482	GCCGAGCAGCGCG	CACTTCAAGGACCTCCGACCGCCTTTGGCAGAGATCAACGGAGATGCT	541
Qy	2172	GGCACAAGCTTCA	CGGTACAGATCAATCAAGATGAGAAATTTCCAGAAAGCTGCAGAACTCAA	2231
Db	542	GGCACAAGCTTCA	CGGTACAGATCAATCAAGATGAGAAATTTCCAGAAAGCTGCAGAACTCAA	601
Qy	2232	GAAAGATTGATTTG	AGCAATGCAATCTTGTGTTCCGGGAAAGGAGTTCACTCGTCTGGG	2291
Db	602	GAAAGATTGATTTG	AGCAATGCAATCTTGTGTTCCGGGAAAGGAGTTCACTCGTCTGGG	661
Qy	2292	CAGCTTACAGCAAG	CTCTCGGGGAAAGGGGCTTCAGACAGCATGTTCTTCTGTTCAGACA	2351
Db	662	CAGCTTACAGCAAG	CTCTCGGGGAAAGGGGCTTCAGACAGCATGTTCTTCTGTTCAGACA	721
Qy	2352	CGTCTCTCTATTA	CAACAGAGCGGGGGCTGACGGCTCCAAATCAGTTTAAAGTCCACGGGCA	2411
Db	722	CGTCTCTCTATTA	CAACAGAGCGGGGGCTGACGGCTCCAAATCAGTTTAAAGTCCACGGGCA	781
Qy	2412	GCTCCGCTCTATG	CAATGCAATGAGAGCAAGCAAGTGGGGGCTGCCCACTG	2471

Db	782	GCTCCCGCTATGACATGATGAGTGAAGAGCGAAGACGAGTGGGGGGTGGCCCACTG	841
Qy	2472	CCTGACCTCTCGGGGGCCAGCGCGAGTCCATCATGTGTGGCCGCCAGTTCTCGGTCCGAGAT	2531
Db	842	CCTGACCTCTCGGGGGCCAGCGCGAGTCCATCATGTGTGGCCGCCAGTTCTCGGTCCGAGAT	901
Qy	2532	GGAGAGTGGGTGTTGAGCATCTCCAGATGGCCATTGACCTGGCGGAGAAAGACAGACGCC	2591
Db	902	GGAGAGTGGGTGTTGAGCATCTCCAGATGGCCATTGACCTGGCGGAGAAAGACAGACGCC	961
Qy	2592	CGCCCTGAGTTCTGTGGCCAGACGGCCCCCTGACAAACAAGTCCCTGTGATGAAGCACCGC	2651
Db	962	CGCCCTGAGTTCTGTGGCCAGACGGCCCCCTGACAAACAAGTCCCTGTGATGAAGCACCGC	1021
Qy	2652	GGCTGACCAAGAGTCAAGAGATGACCTTAAGGGCTCTGGCGCATCTGTGTGAAGCCACGAC	2711
Db	1022	GGCTGACCAAGAGTCAAGAGATGACCTTAAGGGCTCTGGCGCATCTGTGTGAAGCCACGAC	1080
Qy	2712	CCCGGACCGGGGCAACAATGGTGGTCAAGTGGTGGCGGACCGCAACACAGCGTCCAT	2771
Db	1081	CCCGGACCGGGGCAACAATGGTGGTGGCGGACCGGACCGCAACACAGCGTCCAT	1140
Qy	2772	GGTGGACTTCAGCATTCGAGTGGAGAAATCAGTTGTCTGAAAACCTGTGAGAAATTCAA	2831
Db	1141	GGTGGACTTCAGCATTCGAGTGGAGAAATCAGTTGTCTGAAAACCTGTGAGAAATTCAA	1200
Qy	2832	AAACAGCAACGGGTGGCAGAAAGCTGTGGGTGGTGTTCACAAATCTTGCTGTCTTCTTA	2891
Db	1201	AAACAGCAACGGGTGGCAGAAAGCTGTGGGTGGTGTTCACAAATCTTGCTGTCTTCTTA	1260
Qy	2892	CAAAATCAACAGAGACATTCATCCCTTGGCCAGCGCTGTGCTGGCTGCTGCTGCTCAC	2951
Db	1261	CAAAATCAACAGAGACATTCATCCCTTGGCCAGCGCTGTGCTGGCTGCTGCTGCTCAC	1320
Qy	2952	CATCCCTCTGAGTCCGAGAACATCCAGAAAAGACTACGTTTCAAGCTGCACTTCAAGTC	3011
Db	1321	CATCCCTCTGAGTCCGAGAACATTCAGAAAAGACTACGTTTCAAGCTGCACTTCAAGTC	1380
Qy	3012	CCAGCTCTACTACTTCAGAGGCGGAAAGCGAGTACAGTTGCAAAAAGTGGATGAAATGAT	3071
Db	1381	CCAGCTCTACTACTTCAGAGGCGGAAAGCGAGTACAGTTGCAAAAAGTGGATGAAATGAT	1440
Qy	3072	CCGAGTGGCCACACAGCTCTGCTGGCGAACCCACAGTGTGAGCCACAAGAAGTCTCTTGT	3131
Db	1441	CCGAGTGGCCACACAGCTCTGCTGGCGAACCCACAGTGTGAGCCACAAGAAGTCTCTTGT	1500
Qy	3132	GTAT 3135	
Db	1501	GTAT 1504	
RESULT 7			
ADQ84807 ID ADQ84807 standard; cDNA; 3997 BP.			
ADQ84807; 07-OCT-2004 (first entry)			
Human tumour-associated antigenic target (TAT) cDNA sequence #1621.			
human; tumour-associated antigenic target; TAT; cytotoxic; gene therapy;			
cancer; cell proliferative disorder; gene; ss.			
Homo sapiens.			
WO2004060270-A2.			
22-JUL-2004.			
15-OCT-2003; 2003WO-US029126.			
18-OCT-2002; 2002US-0418988P.			

XX (GETH) GENENTECH INC.
PA (WUTD/) WU T D.
PA (ZHOU/) ZHOU Y.
XX
PI Wu TD, Zhou Y;
XX
DR WPI; 2004-534300/51.
XX
PT New nucleic acid molecule and encoded polypeptide, for diagnosing,
PT preventing or treating cell proliferative disorders such as cancer.
XX
DS Claim 1; SEQ ID NO 1621; 5504bp; English.

The present invention describes an isolated tumour-associated antigenic target (TAT) nucleic acid comprising: (a) any of 4622 nucleotide sequences (see SEQ ID NO:1 to 4622); (b) the full-length coding region of (a); (c) the complement of (a) or (b); (d) a sequence that has 80% sequence identity to (a) - (c) or (e) a sequence that hybridises to (a) - (c). Also described: (1) an expression vector comprising the above nucleic acid; (2) a host cell comprising the above expression vector; (3) a process for producing a polypeptide; (4) an isolated polypeptide comprising: (a) an amino acid sequence encoded by any of the above nucleotide sequences; (b) an amino acid sequence encoded by the full-length coding region of the above nucleotide sequences; or (c) a sequence having at least 80% identical to (a) or (b); (5) a chimeric polypeptide comprising the above polypeptide fused to a heterologous polypeptide; (6) an isolated antibody that binds to the above polypeptide; (7) a process for producing the antibody; (8) an isolated oligopeptide that binds to the above polypeptide; (9) a tumour-associated antigenic target (TAT) binding organic molecule that binds to the above polypeptide; (10) a composition of matter comprising the above (chimeric) polypeptide, antibody, oligopeptide or TAT binding organic molecule, in combination with a carrier; (11) an article of manufacture comprising a container and the composition of matter contained within the container; (12) methods and inhibiting the growth of a cell that expresses the above protein, where the growth of the cell is at least in part dependent upon a growth potentiating effect of the above protein; (13) a method of therapeutically treating a mammal having a cancerous tumour comprising cells that express the above protein; (14) a method of determining the presence of a protein in a sample suspected of containing the protein described above; (15) methods of diagnosing the presence of a tumour in a mammal; (16) a method for treating or preventing a cell proliferative disorder associated with increased expression or activity of the above protein; and (17) a method of binding an antibody, oligopeptide or organic molecule to a cell that expresses the protein described above. The TAT sequences have cytostatic activities, and can be used in gene therapy. The composition and methods are useful for diagnosing, preventing or treating cancer. The composition is also used for preparing a medicament for the therapeutic treatment or diagnostic detection of a cell proliferative disorder or cancer. The present sequence represents a human TAT cDNA sequence from the present invention.

SQ Sequence 3997 BP; 998 A; 1132 C; 1068 G; 799 T; 0 U; 0 Other;

Query Match	33.1%;	Score 1037.4;	DB 12;	Length 3997;
-------------	--------	---------------	--------	--------------

Matches 1864; Conservative 0; Mismatches 1186; Indels 45; Gaps 7;

Oy	21	GCCGACCCAGGATACAGACTGGGGGCCCCGAAAATTCGGGGATCAGTACTTTGAAACG	80
Db	150	GCAGACTGCAAGGGATGCGCTTGGGTGCCCAACCCCTGTGGAAATTGACCCCTTGAAGCC	209
Oy	81	TGGACGAAGACGGCCCCCAACCTTCAGGAAAATCTGTCTCATCAAAATCCAGATGCT	149
Db	210	TGGGCGAAGTCTCTTGGCCAGATGCAAGAGAGACCTGACCTCAGAGTAAAGCTGCT	265
Oy	141	GGATGACACCCACGAGAGGCATTTGAAAGTTCACAAAGAGCTCTGGGAAAGGTGCTGTGA	200
Db	270	GGACAAACACATGGAAATATTTGACATTGAGCTTAAATGCGATGGCCAGGTATTACTGAC	329
Oy	201	TGCAGTTTGCAACCACTCAGCTCTGTGGAAAGGTGACATATTTTGGCTCCAGTTTCTGA	260

D	b	330	ACAAGTGTGGAAGCCTTTAAACCTGTGTGAGAATGTGACTACTCTCGGATGTGAGCTTTCAA	385
O	y	261	TCACAAAAGATCACGGTGTGGCTGGATCTCTTAAACCATTGTGAAACAGATTAGA	320
D	b	390	TACTCAGTCTCTACGGATTTGGCTTGAACCTATGAACCATCATTTAGGCAAAATCGA	449
O	y	321	GCCAAAGCACCTGTGTTTGAATTGTGTGAAATTTCTTCGCGCTGACACACAACT	380
D	b	450	GCCAAAGATGTGGTCTTCGCTAGCTGTAAATTTTTCACCTGTATCTGTGACACT	509
O	y	381	CCAAAGAACTCACAGGTACTGTGTGCGCTGCAAGGTGAAGCAGGACTTGGCTCAAG	440
D	b	510	ACAAAGAAATATACAGATCTTTGTTCCTTGCACTTAAGAGAACCTCTGGAAGA	569
O	y	441	CAGTTGACGTGTAAATGACACCAAGCGCACTCTTGAATTCACATTTGTGCAATCTGA	500
D	b	570	GCGTTTGACCTGTGCTGACACCAAGCGGCTTCTCAAGTCCATCTCTGCACTCGGA	629
O	y	501	GATTGGGATTTTATGTAACCTTGGACAGAGAGCATTAAGCAAAAATAATATACATAC	560
D	b	630	AATAGGAATTAACGATGAACCGCTGACGAGAGCAGCTCAAAAGTGAACGAGTATTTC	689
O	y	561	TCAGAAACGCACTAAGAGSACAAATGTGGAATTTACCAATACCATTTGACAAAC	620
D	b	690	TGGCCAGACGACCTGCTTGAAGAAATCTAATAATTCATCAAGAGCAGTGGGCGAC	749
O	y	621	ACCAGCAAAATCAGATTTTCAGGCTCTAGAGATTCGCGCTGAGATGTATGAT	680
D	b	750	ACTGCTGATGTGATTTCCAGGTCTCGAAATCTCGAAAGTTGAAATGTACGGCAT	809
O	y	681	CCGGTTGACCCCGGCGCAAGACAGGGAAGGACGAAATCAATCTGGCCGTGGCAAC	740
D	b	810	CAGATTTCACTAGCTTCTGACAGGGAAGGAACAAAGATTCAATCGGACGTTCCACAT	869
O	y	741	GGGAATTTAGTGTTCAGGCTTTCATTAAGTCAATGCTTCACTGGGCGCAAGTCCG	800
D	b	870	GGGTGTACTCGGTTCAGGGGACACCAAAATCAACTTTCACCTGCTCCAGGTCCG	929
O	y	801	GAACTGAGCTTCAAGAGAAAGCGCTTCTCATCAAGCTCGGCGCAGATGCCAATGTGC	860
D	b	930	TAACTAAGCTTCAAGAGAAAGATTTCTTATCAAACTTCAATCCAGAGTTTACGACC	989
O	y	861	GTACCAGATACCTTGAATTCGTATGCGCAGTGGGATTTCTGCAAGTCTTCTGAA	920
D	b	990	TTACAGAGACATTAAGAATTTTGTGGGTAGATGAATGTAAGAATCTTGGA	1049
O	y	921	AATCTGTGTGAACATCATGCTCTTTAGACTTTTGAAGGCCCAACCAAGCCCA	980
D	b	1050	GATTTGTGTGAGTACACACTTTTTAGACTTTTGAACCACTTAAGCCMAAGCAA	1109
O	y	981	GCCGCTCTCTTAAAGCGGGGGTCAATTCCTCGTTCAGTGGTGGACTCAAGACAGT	1040
D	b	1110	AGCGTCTTCTTCAAGCCGGGGCTCTCTCTTCAGATACGTGGAGAACTCAAGAACT	1166
O	y	1041	TCTGACTATGTAAAGAGAGGACATTAAGAGGTGCAATTGTAAGAAAGACACGAA	1100
D	b	1170	AGTAGATTAATTAAGAGACAGTGGAAATGAAGATTCATATGAAGAAAGGACAGCA	1222
O	y	1101	GATTCATTCTATCCGAGACCTTGTTCACAGCTTACAGAACTGAATTTGGAAGTCTGGA	1166
D	b	1230	GACCCACACGTCC-----GTTGAGACTCTGACTGACACTACCA	1271
O	y	1161	GCAGTCTAGAGAGACACAGCCTTAACATTTGGAAGAAAGTGCAGATCTCCAGGGGGCA	1222
D	b	1272	ACAGAGCATTCATTTCCCGAGAGGATTAAGGACTCTGCTCCCATCTTCAAGCAATGC	1331
O	y	1221	GAGCTGCCGCGAGGAAAGAAACGAAAGTTTCCGCGGGAGGCCGGGATCGACCCGAG	1280
D	b	1332	CTTTTACTCGGTCTCTCCCTCACTCTGCTCCCTCTGCGCGAGATTTAAAGACAG	1391
O	y	1281	CCCTGCGCGAGAGAGAGCCCGCGGGTAAACAAGAGGCGAGCGGCGCTCGGCGCC	1340
D	b	1392	CAGCAGCTCTCTCAAGATCTCCCAAGTGTCTTCACTCAAGAGTCCAGCTGACAGAGGCG	1451

OY	1341	CAACGAGGAAGAAGAGAG---TCGTTAAGATATAGACCCAGACAGATTAACCTCAGCC	1339
Db	1452	CAGTGGACAAATGGCTGGAGGCCCCGACACCAATCGGCCACAGCCCTCGGGCCCCCGC	1511
OY	1398	CCCCAGCCCAAGCACAGAGCTCCTGTAGCTGGAGATCTTCAACCTTTCCAGAGCTGTCTGTAA	1457
Db	1512	ACTCCAGCTGTGTCCAGGCCCTTTCCACGAAAGATCTCAGCTCTTCCCTCCAGCCGGAA	1571
OY	1458	CTCCGAGGGGGAGTGGCCCTTGCCAACTGACCTTGTCTCCAACTGAGCCCCGACAC	1517
Db	1572	GAGCCCCCTGAGTGTGAGCCCTCGAATTCACAGTGCCTT-----TGGCCCCAGCTGA	1622
OY	1518	CAAGCAGGCCCTCTCCCTTGATCAGCCCGCTGTGAATGACAGGCCCTGCCCCGGACGGA	1577
Db	1623	ACAGGGCTCATCCCACTCTTGAGCCCTGTCTTCACTGATCTGTGGCGGACCGGGATGGA	1688
OY	1578	CGATGAGGATGAGGGCCGGAGGAAGATTCACCACTGATTAAGCGTACTTCATAGCTTA	1637
Db	1683	CTGCGAG---GAGCCACAGACAAAGCGGTGCTGTGACAGACGAGGCGTACTTCATAGTCAA	1739
OY	1638	GGAAGTGTCTACCAACGAGCGGAACATATCTGAAGATCTCGAAGTATCACTTGTGTGTT	1697
Db	1740	AGAAATTCGCTCTCAGAAACGAACATACCTCAAGGATTTAAGATTAATTAACGTTGAGTT	1799
OY	1698	TCAGAGCACAATGAGCAAAAGAGACGCCATGCCGGAAGACCTGAAAGTCTCATATTTCC	1757
Db	1800	CCGCAACGCCAGTGGTGAAGAGAGACGCCATGCTGTGCACTTGAATGACGCTGCTTTCTC	1855
OY	1758	GAAATTTGAACTTTGCACAAAATTTCACTAATTTTCTCAAGAAATTTGACACGACT	1817
Db	1860	CAATCATGATCCCACTATAGATTCCACAGAGGCTCTCGCCAGGAGTGAGCAGAGGCT	1919
OY	1818	TGCCCTGGGGAAGGCGGCTCAAAATGCCCAATCAGAG---ATTACCAAAATACGGGGA	1874
Db	1920	GGCACTCTGGAAAGGCCCCCTCCAAAGCCACAAAGGACGTCACTCAAGAAATCGGGGA	1979
OY	1875	TGTCAATGCTGAAGAACTTCAGGGCATAAGACACTGGCGGCTCACTGTGAAACACAG	1934
Db	1980	CATCTGCTCAGAAACATGCGCCAGTTAAAGAGATTAAACAGTACTTCCAAACATAGA	2039
OY	1935	CGAAGCCTTGGAGGCCCTCTGAGGAATGGAATCAAGATCCCGGGGCTGAGGAATCTGTG	1994
Db	2040	CGAGGTCTTAACAACACTGGAAAAAGGCTAACMAACGCTGAAGAAATTGGAGGACGTETA	2099
OY	1995	CAGAGACTTTGAGCTGAGAGAGTGTGTATCTACCGCTCAACACTTCTCTGCGGCC	2054
Db	2100	CAAGGAGTTTGAAGCTGACGAAGGTCTGTACTTGCCTCAACAGTTTCTGCTGAAGCC	2159
OY	2055	ACTGCACCGGCTCATGCACTACAAAGCAAGTCTTGAAGCGCTGTGCAACACACCCGCC	2114
Db	2160	CATCCAGGGGCTGTGCACTACCCCTGCTGCTCAGCGCTTAAGCGGACATTCAGGCC	2219
OY	2115	GAGCACGCGGCACTTCAAGGGAATGCGCGCGCTTTGGCAGAGATCAACGAGATGGTGGC	2174
Db	2220	CGGGCAACATGACTACGCTGACTCCTCAATGACGCCCTTAAGACCATCACAGAGGTGACAC	2279
OY	2175	ACAGCTCCACGGTACGATGATCAAGATGAGAAATTTCCAGAAAGCTGACGAACTCAAGAA	2234
Db	2280	CACACTACAGACATTTCTCATTCGCGCTGGAGAAACCTGCAAGAAAGCTTAACGAGCTGACGCG	2339
OY	2235	AGATTTGATTTGGCATTTGACAACTTTTGTGTCTCCGGAAAGGAATTCATCCGTCTGGCAG	2294
Db	2340	GGACCTGGTGGCAATTAAGAACTTCATTTGCTCTGGCAGGGAGTTCAATCCGTGAGGGCTG	2399
OY	2295	CCTCAGCAAGCTCTCGGGGAAAGGGGCTCCAGACACGCAATGTTCTTCTGTTCACAGAGT	2354
Db	2400	CTTTCACAGCTCAACAAAGAGGCCCTGACAGCAAGGATGTTTTTCTGTTCTCAAGTAT	2459
OY	2355	CTGTCTATACAGAGCGGGGGCTGACGCGCTCCAATCAATTTAAAGTCCACGGGACGT	2414
Db	2460	GTTCCTGTATCACAGCAAGAGAGTTGACGGGACACAGCACTTCCGAAATCCGGGGCTCTCT	2519

QY 2415 CCCGCTCTATGGCATGACATTTGAGAGAGCGAAGACGATGGGGGGTGGCCCCACTGGCTT 2474
DB 2520 TCCCTCTCAAGGCAATCTGTGTGAAGAAAGTGAATACGATGTCTCTTCCACATGTTT 2579
QY 2475 GACCCCTCCGGGGCCAGCGGAGTCATCATCTGTGGCCGCAATTTCTCGATCCGAGATGA 2534
DB 2580 CACCATCTACGGGGCTCAGAAAACAATCGTGTGGCAGCAGACTCGGCTGGAGAAAAGA 2639
QY 2535 GAATGGGTGAGAGACATCCAGATGGCCATTTGACTGTGGGGAGAAAAGACAGACCCCGC 2594
DB 2640 GAATGTGAATGCTGTGAACCTGACCTCCGATCCAAAGCAGCAAGATGGCGGTGACACG 2699
QY 2595 CCCTGATTCCTGTGGCAGCA-----GCCCGCCCTGCAACAAGTCCCTGATGAAGCCAC 2648
DB 2700 CCTGTACGTCCAGCGCCGCACTGTGTCACTGTCCCCCAAGATCCCCCAAGAGATATC 2759
QY 2649 CCGGGCTGACAGAGATCAGAGATGACCTGAGCGCTTCGCGCACTGTGTGAGCGCA 2708
DB 2760 TCTG---GAGCAGAGATCAGAAAGATGATGCTCGGGGTGTCCGAGCTCCCTGGAGGGGCA 2816
QY 2709 GGGCCCGGACCGCGGCAACATATGTTGACCTGTGTGACCCGCAACAGCAGCTTC 2768
DB 2817 TGGCCAGCACCGGGCCAAACCAATGACGTGTGTGTACCGGAACACGCGTGTTC 2876
QY 2769 CATGTGACTTCAGCATCGCAGTGTGAGAAATGTTGTGAAACCTGTGAGAAAT 2828
DB 2877 CAGGGACAGACACAGTGCAGCTGTCCAGAACAGCTTTACAGATATCTGTGAAGAAAT 2936
QY 2829 CAAAAACAGCAACGGGTGCGAAGCTGTGGGTGTGTTCACAACTTGTGCTGTTCTT 2888
DB 2937 CAAAAACATGCTGTGCGCAGAAAGCTGTGGGTGTGTTCCTTTACCACTTGTGTGTTCT 2996
QY 2889 CTAACAAATACACACAGACATATATCCCTTGGCAACCTGCGCTGTGCTGCTGCT 2948
DB 2997 CTAACAAATCATCAGATGATGATACCCACCTGCGCAGCTTCCTGTGTGTGTGTGT 3056
QY 2949 CACCATCCCTCTGATGCTCGAGAACATCCAGAAAGACTAGTGTCAAGCTGCACTTCAA 3008
DB 3057 GAGCATCCCCAGGGAAGCCGATGGCATACCAAAAGCTATGTTTGAAGCTCCAGTTCAA 3116
QY 3009 GTCCACGTTCTACTTCTCAAGGGCGGAAAGCAAGTACAGTTTGAAGGTGATGAGAT 3068
DB 3117 ATCCCACTCTACTTCTTCCGGGCTGAGAGCAAGTACATTTGAAGGTGATGAGAT 3176
QY 3069 GATCCGAGTCCACAGCTGTGCTGCGGAGCC 3103
DB 3177 GATCCAGGGGGCCAGAGCTCAGCCGGGAGGGCC 3211

RESULT 8
ADR25675 standard; DNA: 3997 BP.
XX ADR25675;
XX
DT 21-OCT-2004 (first entry)
XX
DE Breast cancer prognosis marker #1536.
XX
KW de; breast cancer; prognosis; gene expression; diagnosis.
XX
OS Homo sapiens.
XX
PN WO2004065545-A2.
XX
PD 05-AUG-2004.
XX
PF 15-JAN-2004; 2004WO-US001100.
XX
PR 15-JAN-2003; 2003US-00342887.
XX
PA (ROSE-) ROSETTA INPHARMATICS LLC.
PA (NECA-) NETHERLANDS CANCER INST.

XX
PI Van't Veer LJ, He Y;
XX
DR WPI; 2004-593473/57.
XX
PT Classifying a breast cancer patient according to prognosis comprises
PT determining the similarity between the level of expression of each of
PT five genes in a cell sample taken from patient, to control levels.
XX
PS Disclosure; SEQ ID NO 1536; 226bp; English.
XX
CC The invention relates to a method of classifying a breast cancer patient
CC according to prognosis by determining the similarity between the level of
CC expression of each of five genes for which markers are listed in the
CC specification, in a cell sample taken from the breast cancer patient, to
CC control levels of expression for each respective five genes to obtain a
CC patient similarity value. The methods are useful for classifying a breast
CC cancer patient according to prognosis. Kits and computer program products
CC are useful for data analysis using the diagnostic, prognostic and
CC statistical methods of the invention. This sequence corresponds to a
CC marker used in the method of the invention.

Sequence 3997 BP; 998 A; 1132 C; 1068 G; 799 T; 0 U; 0 Other;
XX
SQ
Query Match 33.1%; Score 1037.4; DB 13; Length 3997;
Best local Similarity 60.2%; Pred. No. 9.5e-267;
Matches 1064; Conservative 0; Mismatches 1186; Indels 45; Gaps 7;
QY 21 GCCGACCCGAGATACGACTGGGGGCCCCGGAATAATCGGGATCATCTTGGAGC 80
DB 150 GCAGCTCAGAGATGCGCTTGGGTGCCAGACCCCTGTGGAGTTACACCCCTGAGCC 209
QY 81 TGGACAGAAAGCGCCGCCAACCTTCAGAAAACTCGTGCATCAAAATCCAGATGCT 140
DB 150 TGGGCAAGCTCTTGGCCAGATGCAAGAGACACTGCACTCAAGTAAAGCTGCT 269
QY 210 TGGGCAAGCTCTTGGCCAGATGCAAGAGACACTGCACTCAAGTAAAGCTGCT 269
DB 141 GGATGACACCCAGAGGATTTGAAGTTCACAAAGAGCTCCTGGAAAGTGTCTGGA 200
QY 270 GGAACAACACATGAAATATTTGACATTGAGCTAAATGGCATGGCCAGATTTACTGAC 329
DB 201 TGCAGTTGCAACCACTCAACTCTGTGAAAGTCAATATTTGGCTTCGAGTTTCTGA 260
QY 330 ACAAGTGTGGAAGCCTTTTAACTGTGTAAATGTACATCTTCGGGATGAGATTTCAAA 389
DB 261 TCACAAAAGATCAGGTGTGGCTGATCTCTTAACCCATGTGAAACAGATTAGAG 320
QY 390 TACTGATCTACTGGAATTTGGCTTGAACCTATGAAACCATCATTTAGCAATTCGAG 449
DB 321 GCCAAGCAGTGTGTGTTAAGTTGTGTGTAATTTCTTCGCTGACCAACAACACT 380
QY 450 GCCAAGAAATGTGTGCTTGGCTGCTGTAATAATTTTCCACTGATCTGTGAGCT 509
DB 381 CCAAGAAATCTCAGAGTACTTGTGCGCTGAGGTGAAGCAGACTTGGCTCAGG 440
QY 510 ACAGAAAGATATACAAATCTTGTGCTTGCATCTTAAGAGACCTGTGGAAGA 569
DB 441 CAGGTTGACGTATGACACACAGGAGCTCTCTGATTTCAACATTTGCAATTTGA 500
QY 570 GCGTTTACCTGTGTGACACACAGCGCCCTTTCAAGTCTTCCTCTGCACTGGA 629
DB 501 GATTGGGATTTTGAAGAGCTTGTGACAGAGACATTTAGCAAAAAATTAATACATAC 560
QY 630 AATAGAGATTTACATGAAAGCTGTGACGAGAGCACTCAAAAGTGAAGATATTGGC 689
DB 561 TCAGCAAGACGACTAGAGAGCAAAATGTGGAATTTCAACCTAACCATTTGACAAAC 620
QY 690 TGGCCAGCAGCACTGCTTGAAGAAATCTAGAAATTCATCAGAGCAGTGGGCGAGC 749
DB 621 ACCAGAGATCAATTTCCAGCTCTAGAGATTGCGGTGGCTAGAGATATGATGAT 680
QY 750 ACTGTGATGTGATTTCCAGGTCTCGAAATTTCTGAAAGTTGAAATGTAGGCAAT 809
DB 681 CCGATTGACCCGGCCAAAGACAGGAAAGCAAGAAATCATCTGCGCTTGCACAAAC 740

XX Open reading frame detection; genome sequencing; colon cancer;
 KW breast cancer; population genome analysis; genetic shift; cancer;
 KM antibiotic resistance; antibiotic non-tolerance; congenital disease;
 KM agriculture; food crop genome; resistance gene; retrovirus;
 KM influenza virus; eukaryotic pathogen detection; trypanosome; Plasmodium;
 KW gene; ss.
 XX
 OS Homo sapiens.
 PN US2002155438-A1.
 EN
 XX 24-OCT-2002.
 PD
 XX 27-SEP-1999; 99US-00406117.
 PF
 XX 20-NOV-1998; 98US-00196716.
 PR
 XX (SIMP/) SIMPSON A J G.
 PA (NETO/) NETO E D.
 PA (BREN/) BRENTANI R R.
 XX
 PI Simpson AJG, Neto ED, Brentani RR;
 DR WPI; 2003-182626/18.
 XX
 XX Determining open reading frames of genome of an organism e.g. a human
 PT suffering from cancer involves use of single oligonucleotide primer at
 PT low stringency for preparing single-stranded cDNA from mRNA of
 PT individual.
 XX
 PS Example 9; Page 105; 959pp; English.
 XX
 XX The invention describes a method of determining open reading frames in
 CC the genome of organism, comprising contacting mRNA from cell of organism
 CC with a single oligonucleotide primer (1) at low stringency, preparing
 CC single-stranded cDNA by reverse transcribing mRNA with (1), amplifying
 CC cDNA, sequencing the product, and repeating the contacting, preparing
 CC and amplifying steps with different primers and sequencing resulting
 CC nucleic acids. The method is useful for: determining that a known
 CC nucleotide sequence from a genome of an organism corresponds to a
 CC nucleotide sequence of an open reading frame; for preparing a contig,
 CC nucleic acid molecule from a genome of an organism; and for sequencing
 CC all or part of a genome of an organism. mRNA is obtained from mammalian
 CC or human cell which is associated with a pathological condition e.g. a
 CC colon cancer or breast cancer cell. The method is useful for analyses of
 CC populations of subjects and can be used to carry out genetic analyses of
 CC large or small populations. Further, it can be used to study living
 CC systems to determine if, e.g. there have been genetic shifts which render
 CC an individual or population more or less likely to be afflicted with
 CC diseases such as cancer, to determine antibiotic resistance or non-
 CC tolerance, and so forth. The method can also be used in the study of
 CC congenital diseases, and the risk of affliction to a fetus, as well as
 CC the study of whether the conditions are likely to be passed to offspring
 CC through ova or sperm. The analyses for pathological conditions can be
 CC carried out in all animals, plants, birds, fish, etc. Using this method,
 CC in the area of agriculture, for example the genomes of food crops can be
 CC studied to determine if resistance genes are present, defects in plant
 CC genomes can also be studied in this way. Similarly, the method permits
 CC determination of the pathogens which integrate into the genome, such as
 CC retroviruses and other integrating viruses such as influenza virus, have
 CC undergone shifts or mutations, which may require different approaches to
 CC therapy. This method is also applied to eukaryotic pathogens, such as
 CC trypanosomes, different types of Plasmodium, etc. The method essentially
 CC eliminates sequencing of non-coding portions. This sequence represents a
 CC polynucleotide isolated from human colon cancer cell cDNA library
 CC
 XX
 SQ Sequence 582 BP; 109 A; 147 C; 185 G; 140 T; 0 U; 1 Other;

Query Match 17.6%; Score 550.4; DB 10; Length 582;
 Best Local Similarity 99.6%; Pred. No. 1.3e-136;
 Matches 551; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2583 CAGCAGCCCCCGCCCTGAGTTCCTGCGCAGCAGCCCCCTGACAAAGTCCCTGATGA 2642
 DB |||||||
 QY 571 CAGCAGCCCCCGCCCTGAGTTCCTGCGCAGCAGCCCCCTGACAAAGTCCCTGATGA 512
 DB |||||||
 QY 2643 AGCCACCGCGGCTGACCGAGATCAGAGATGACTGAGCCGCTTGCGCAATCGCTGGA 2702
 DB |||||||
 QY 511 AGCCACCGCGGCTGACCGAGATCAGAGATGACTGAGCCGCTTGCGCAATCGCTGGA 452
 DB |||||||
 QY 2703 GCGCAGGCGCCCGCAGCGCGGCAACAGATGGTGCAGCTGTGCTGAGCCGCAACCCAG 2762
 DB |||||||
 QY 451 GCGCAGGCGCCCGCAGCGCGGCAACAGATGGTGCAGCTGTGCTGAGCCGCAACCCAG 392
 DB |||||||
 QY 2763 CGTCTCAATGATGACTTCAAGATGCAAGTGAAGATGATGTTGTGAGAACTGCTGAG 2822
 DB |||||||
 QY 391 CGTCTCAATGATGACTTCAAGATGCAAGTGAAGATGATGTTGTGAGAACTGCTGAG 332
 DB |||||||
 QY 2823 GAAATTCAAAAACAGCAACGGGTGGCAGAAAGTGTGGGTGTTCACAAAATTTCTGCT 2882
 DB |||||||
 QY 331 GAAATTCAAAAACAGCAACGGGTGGCAGAAAGTGTGGGTGTTCACAAAATTTCTGCT 272
 DB |||||||
 QY 2883 GTTCTTCTACAAATACACACAGGACATATCATCCCTGCGCAGCTGCTCTGCGCTA 2942
 DB |||||||
 QY 271 GTTCTTCTACAAATACACACAGGACATATCATCCCTGCGCAGCTGCTCTGCGCTA 212
 DB |||||||
 QY 2943 CTGCTCAACATCCCTCTGAGTCCGAGAACATCCAGAAAGCTACGTGTTCAAGCTGCA 3002
 DB |||||||
 QY 211 CTGCTCAACATCCCTCTGAGTCCGAGAACATCCAGAAAGCTACGTGTTCAAGCTGCA 152
 DB |||||||
 QY 3003 CTTCAAGTCCCAAGCTCTACTTCAAGGCGGAAAGCGAGTCAAGTTGGAAGGTGAT 3062
 DB |||||||
 QY 151 CTTCAAGTCCCAAGCTCTACTTCAAGGCGGAAAGCGAGTCAAGTTGGAAGGTGAT 92
 DB |||||||
 QY 3063 GGAAGTGAATCCGACAGTGCACAGAGCTGCTGCGGACCCGACGTTGAGCCACAAAGA 3122
 DB |||||||
 QY 91 GGAAGTGAATCCGACAGTGCACAGAGCTGCTGCGGACCCGACGTTGAGCCACAAAGA 32
 DB |||||||
 QY 3123 GTCTCTGTGTAT 3135
 DB |||||||
 QY 31 GTCTCTGTGTAT 19
 DB |||||||

RESULT 11
 ACD94832/C
 ID ACD94832 standard; cDNA; 582 BP.
 XX
 AC ACD94832;
 XX
 DT 23-SEP-2003 (first entry)
 XX
 DE Human colon cancer cell expressed cDNA #3244.
 XX
 KW Open reading frame detection; genome sequencing; colon cancer;
 KW breast cancer; population genome analysis; genetic shift; cancer;
 KM antibiotic resistance; antibiotic non-tolerance; congenital disease;
 KM agriculture; food crop genome; resistance gene; retrovirus;
 KM influenza virus; eukaryotic pathogen detection; trypanosome; Plasmodium;
 KW gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN US2002155438-A1.
 EN
 XX 24-OCT-2002.
 PD
 XX 27-SEP-1999; 99US-00406117.
 PF
 XX 20-NOV-1998; 98US-00196716.
 PR
 XX (SIMP/) SIMPSON A J G.
 PA (NETO/) NETO E D.
 PA (BREN/) BRENTANI R R.
 XX
 PI Simpson AJG, Neto ED, Brentani RR;

```
XX WPI; 2003-182626/18.
DR
XX
XX Determining open reading frames of genome of an organism e.g. a human
PT suffering from cancer involves use of single oligonucleotide primer at
PT low stringency for preparing single-stranded cDNA from mRNA of
XX individual.
XX
XX Example 9; Page 477; 959pp; English.
XX
XX The invention describes a method of determining open reading frames in
CC the genome of organism, comprising contacting mRNA from cell of organism
CC with a single oligonucleotide primer (I) at low stringency, preparing
CC single-stranded cDNA by reverse transcribing mRNA with (II), amplifying
CC cDNA, sequencing the product, and repeating the contacting, preparing
CC and amplifying steps with different primers and sequencing, resulting
CC nucleic acids. The method is useful for: determining that a known
CC nucleotide sequence from a genome of an organism corresponds to a
CC nucleic acid molecule of an open reading frame; for preparing a contig;
CC all or part of a genome of an organism. mRNA is obtained from mammalian
CC or human cell which is associated with a pathological condition e.g. a
CC colon cancer or breast cancer cell. The method is useful for analyses of
CC populations of subjects and can be used to carry out genetic analyses of
CC large or small populations. Further, it can be used to study living
CC systems to determine if, e.g. there have been genetic shifts which render
CC an individual or population more or less likely to be afflicted with
CC diseases such as cancer, to determine antibiotic resistance or non-
CC tolerance, and so forth. The method can also be used in the study of
CC congenital diseases, and the risk of affliction to a fetus, as well as
CC the study of whether the conditions are likely to be passed to offspring
CC through ova or sperm. The analyses for pathological conditions can be
CC carried out in all animals, plants, birds, fish, etc. Using this method,
CC in the area of agriculture, for example the genomes of food crops can be
CC studied to determine if resistance genes are present, defects in plant
CC genomes can also be studied in this way. Similarly, the method permits
CC determination of the pathogens which integrate into the genome, such as
CC retroviruses and other integrating viruses such as influenza virus, have
CC undergone shifts or mutations, which may require different approaches to
CC therapy. This method is also applied to eukaryotic pathogens, such as
CC trypanosomes, different types of Plasmodium, etc. The method essentially
CC eliminates sequencing of non-coding portions. This sequence represents a
CC polynucleotide isolated from human colon cancer cell cDNA library
XX
XX Sequence 582 BP; 109 A; 147 C; 185 G; 140 T; 0 U; 1 Other:
SQ
Query Match 17.6%; Score 550.4; DB 10; Length 582;
Best Local Similarity 99.6%; Pred. No. 1.3e-136;
Matches 551; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 2583 CAGCAGCCCCCGCTGAGTTCCTGCGCAGCAGCCCCCTGACACAAAGTCCCTGATGA 2642
DB 571 CAGCAGCCCCCGCTGAGTTCCTGCGCAGCAGCCCCCTGACACAAAGTCCCTGATGA 512
QY 2643 AGCCACCGCGGCTGACAGAGATCAGAGATGACCTGAGGCGCTGCGCAGCAATCGGTGA 2702
DB 511 AGCCACCGCGGCTGACAGAGATCAGAGATGACCTGAGGCGCTGCGCAGCAATCGGTGA 452
QY 2703 GCGCCAGGCCCCGACCGCGGCAACAAATGTGTGCTGTGACCGCAACACAG 2762
DB 451 GCGCCAGGCCCCGACCGCGGCAACAAATGTGTGCTGTGACCGCAACACAG 392
QY 2763 CGTCTTCATGTGTGATTCAGCATTCGAGTGAAGATCAGTTGTCTGAAACCTGCTGAG 2822
DB 391 CGTCTTCATGTGTGATTCAGCATTCGAGTGAAGATCAGTTGTCTGAAACCTGCTGAG 332
QY 2823 GAAATTCAAAACAGCAACGCGGTGCGAAGCTGTGAGGTTCACAAACTTCGCGCT 2882
DB 331 GAAATTCAAAACAGCAACGCGGTGCGAAGCTGTGAGGTTCACAAACTTCGCGCT 272
QY 2883 GTTCTTCTACAAATCAGCAGCAGCAATCATCCCTTTCAGGCTGCTGTCTGCGCTA 2942
DB 271 GTTCTTCTACAAATCAGCAGCAGCAATCATCCCTTTCAGGCTGCTGTCTGCGCTA 212
```

```
QY 2943 CTGCTCACCATTCCCTCTGAGTCCGAGAACATCCAGAAAGACTAGCTGTCACTGCA 3002
DB 211 CTGCTCACCATTCCCTCTGAGTCCGAGAACATCCAGAAAGACTAGCTGTCACTGCA 152
QY 3003 CTTCAGTCCCAAGCTTACTTACTTTCAGGCGGAAACGAGTACACGTTGGAAGTGAT 3062
DB 151 CTTCAGTCCCAAGCTTACTTACTTTCAGGCGGAAACGAGTACACGTTGGAAGTGAT 92
QY 3063 GGAAGTATCCGAGTGCACACAGCTCTGCTGCGACCCACAGTGTGAGCCACAAGA 3122
DB 91 GGAAGTATCCGAGTGCACACAGCTCTGCTGCGACCCACAGTGTGAGCCACAAGA 32
QY 3123 GTTCTTCTGCTAT 3135
DB 31 GTTCTTCTGCTAT 19
RESULT 12
AD063863
ID AD063863 standard; cDNA, 3200 BP.
XX
XX AD063863;
AC
XX
XX 07-OCT-2004 (first entry)
DT
XX
XX Novel human cDNA sequence #1024.
DE
XX
XX ss; gene; osteopathic; neuroprotective; nootropic; antiparkinsonian;
KM cytostatic; gene therapy; diagnostic marker; morbid state; osteoporosis;
KM neurological disease; Alzheimer's disease; Parkinson's disease; dementia;
XX cancer.
XX
XX Homo sapiens.
OS
XX
XX EP1440981-A2.
PN
XX
XX 28-UTL-2004.
PD
XX
XX 21-JAN-2004; 2004EP-00001196.
PF
XX
XX 21-JAN-2003; 2003JP-00102206.
PR
XX
XX 09-MAY-2003; 2003JP-00131392.
XX
XX (REAS-) RES ASSOC BIOTECHNOLOGY.
PA
XX
XX Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
PI Yamamoto J, Isono Y, Nagai K, Irie R;
PT P-PSDB; AD066051.
DR
XX
XX WPI; 2004-535376/52.
DR
XX
XX Novel 2495 cDNA, useful for treating osteoporosis, neurological diseases,
PT Alzheimer's diseases, Parkinson's diseases, dementia and various cancers.
PT
XX
XX Claim 1; SEQ ID NO 1024; 2449pp; English.
PS
XX
XX The invention relates to 2495 novel polynucleotides (I) and their encoded
CC polypeptides, sequences hybridizing to these nucleotides, sequences
CC encoding partial polypeptides and sequences having 70% or 90% identity to
CC the nucleotide and protein sequences. The nucleotides and polypeptides
CC are useful as diagnostic markers or therapeutic target for the diseases
CC or morbid states. They are also useful for treating osteoporosis,
CC neurological diseases, Alzheimer's diseases, Parkinson's diseases,
CC dementia and various cancers. This sequence corresponds to a nucleotide
CC sequence of the invention.
XX
XX Sequence 3200 BP; 951 A; 659 C; 718 G; 872 T; 0 U; 0 Other:
SQ
Query Match 14.3%; Score 447; DB 12; Length 3200;
Best Local Similarity 65.5%; Pred. No. 1.6e-108;
Matches 654; Conservative 0; Mismatches 345; Indels 0; Gaps 0;
```



```
OY 96 CCCAACACCTTGAGGAAAACTCGTGCATCAAAATCCAGATGCTGATGACACCCAGGA 155
DB 160 CCCAACGGCTCGCTGATGATGCTACATTTAAAGGACGTTTGTGATGATTTCCAGAA 219
OY 156 GGCATTGTAAGTTCCACAAAGAGCTCTCGGAAAGGTGCTGCTGATGACGTTTGCACCA 215
DB 220 GATTTTGTGTTGATCAAAAGATCATCCGGAAAGGATGTTTAACTGAGTTGCGCCA 279
OY 216 CCTCAACCTCGTGGAGGTGACTACTATTTGGCCCTCGAGTTTCCGATCACAAAAGATCAG 275
DB 280 TCTAAATCTTCTGAAAAGAAATATTTGGATTGAATCTGACGCTATTCGAAATATA 339
OY 276 GGTGTGCTGATCTCTCTAAACCCATTGTGAACAGATTAGAGGCCAAAGACGTTGT 335
DB 340 TGTGTGCTGAGACTTTGAGCCCATTAACAAAGAGGTAAABAAATCTTAAGAGATTGT 399
OY 336 TGTTAAGTTTGTGTGAAATTTCTTTCCGCTGACCAACACAACTCAAGAAAGACTCAC 395
DB 400 TTTCAATTATGTGTGAAATTTTCCAGTGAACCTGAGACATCTGCGGAAAGAACTTAC 459
OY 396 AAGGTACTGTTCCGCTGAGGTGAAGACGACTGGCTCAAGGCGAGTTGACGTTAA 455
DB 460 AAGGTACTTTTACTCTTCAATTAAGAAAGATTTGGCTCTAGGAAGGCTTCATGACG 519
OY 456 TGAACACGACGAGCTCTCTGATTTACACATTTGCAATCTGAGATTGGGATTTTGA 515
DB 520 TGAACACTGTAACGCGTTGATGTATCTACATCTTACATGAGAACTTGGAGACTTTCA 579
OY 516 TGAAGCTTGGACAGAGACCTTAGCAAAAATAATACATCTCTACAGACAGCACT 575
DB 580 TGAAGAAACAGATGAGAACATCTGGACAAACTCGGTCTTAAACCAACGACGTTT 639
OY 576 AGAGCAAAATCGTGAATTTTCAACCATTAACCATTTGACAAACCCAGAAATCAGA 635
DB 640 AGAGGCAAGATCATGACTTTTCATAGAAAGCAATTTGGAGAGAGCCAGCTGAATCTGA 699
OY 636 TTTCCAGCTCTAGAGATTTCCCGTGGCTAGAGATGTATGAAATCCGTTGACCCGCG 695
DB 700 CATTTGCTACTGACATAGCAAGAAAGCTGATGTATGTATGGATCAGGCTCACCCTGC 759
OY 696 CAAGACAGGAAAGGACAGAAATCAATTTGGCCCTTGGCAACAGGAAATCTAGTGT 755
DB 760 CAGTATGTGTGAAGGATGACAAATTCACCTGGTGTGTCTCACTGGAGTACTGTGT 819
OY 756 TCAGGGTTTCACTAATGATCAATGCTTCAACTGGGCAAGTGGCGAAGCTGAGCTTCA 815
DB 820 ACGGGAAATACAAAGATCAATCTTTAACTGGCTAAATCCGAAAGTTGAGTTTAA 879
OY 816 GAGAAAGCGCTTTCTCATCAAGCTCCGACCAATGCTCAATGCTGACAGATACCTT 875
DB 880 GAGAAAGCAATTTCTCATCAAACTTATGCTCAATATCTTGTGTGTGCAAGATACCTT 939
OY 876 GGAATTCCTGATGCGCAGTGGGATTTCTGCAAGTCTTCTGGAATCTGTGTGAACA 935
DB 940 GGAGTTCAACATGCGCAGCGAGATGCTGCAAGGCTTTCTGGAACACTGTGTGAATA 999
OY 936 TCATGCTCTTTTAAAGCTTTTGAAGAGCCCAAAACCAAGCCCAAGCCGCTCTTTAG 995
DB 1000 CATGCTTTCTTCAGGCTTTGGAAGAGCCCAATTAAGCCCAAAACCTTACTCTGCG 1059
OY 996 CCGGGGCTCATATTTCCGTTAGTGTGAGTCTGAGACTCAAGAGAGGTTCTGACTATGTTAA 1055
DB 1060 CAAGGGTTCAAGTTTCGCTATATGTGAGAGAACCCAAAGGCAACTTTGGAATATGTGAG 1119
OY 1056 AGAAGAGACATTAAGAGTGTGAGTTGAAAGGAACA 1094
DB 1120 AAAAGGAGGCTGAAAGACTTGCATTTGAAGGAACA 1158
```

```
RESULT 13
AC161462
ID AC161462 strand; cDNA; 435 BP.
XX
```

```
AC AC161462;
XX
DT 24-MAR-2005 (first entry)
XX
DE Human colon cancer differentially expressed polynucleotide, SEQ ID:7597.
XX
KW Differential expression; diagnosis; therapy; drug screening; cancer;
KW neoplasm; colon tumor; breast tumor; pancreas tumor; cytostatic; vaccine;
KW ss.
XX
OS Homo sapiens.
XX
PN WO200500087-A2.
XX
PD 06-JAN-2005.
XX
PF 13-MAY-2004; 2004WO-US015421.
XX
PR 03-JUN-2003; 2003US-0475872P.
XX
PA (CHIR ) CHIRON CORP.
XX
PI Randazzo F, Moler E, Escobedo J, Garcia PD;
XX
DR WPI; 2005-075421/08.
XX
PT New isolated polynucleotides, which are differentially expressed in colon
PT cancer cell, useful for treating cancer, e.g. colon cancer, breast
PT cancer, or pancreatic cancer.
XX
XX
XX Claim 1; SEQ ID NO 7597; 97bp; English.
XX
XX
CC The invention relates to 9672 polynucleotides (AC15866-AC161537) which
CC are differentially expressed in colon cancer cells. The invention also
CC relates to vectors and host cells comprising a differentially expressed
CC polynucleotide of the invention; a method for detecting a cancerous cell
CC by detection of a gene product of the polynucleotides; a method for
CC inhibiting a cancerous phenotype of a cell by inhibiting a gene product
CC of the polynucleotides; a method of treating an individual with cancer by
CC administration of a modulator of a gene product of the polynucleotides;
CC and an isolated antibody that specifically binds to a polypeptide encoded
CC by one of the 9672 polynucleotides. The polynucleotides, polypeptides,
CC antibodies, and methods are useful for the detection of cancerous cells;
CC for the diagnosis, prognosis and management of cancer; for the
CC identification of agents that modulate the phenotype of cancerous cells;
CC and for the identification of therapeutic targets for cancer chemotherapy;
CC and for the treatment of cancer, especially colon cancer and metastasized
CC colon cancer, but also breast or pancreatic cancer. The polynucleotides
CC are also useful as a source of probes or primers for use in diagnostic
CC methods. The differentially expressed polynucleotides or their encoded
CC proteins can additionally be used as vaccines to modulate primary immune
CC responses for the prevention or treatment of cancer. The present sequence
CC represents a specifically claimed polynucleotide which is differentially
CC expressed in colon cancer. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 435 BP; 100 A; 142 C; 115 G; 78 T; 0 U; 0 Other;
```

Query Match 13.4%; Score 420.8; DB 14; Length 435;
Best Local Similarity 99.3%; Pred. No. 6.1e-102;
Matches 433; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

```
OY 2540 GGGTTGAGAGACATCCAGATGGCCATTGACCTGGCGGAAAGAGCGAGCCCGGCCCTTG 2599
DB 1 GGGTTGAGAGACATCCAGA-GGCCATTGACCTGGCGGAAAGAGCGAGCCCGGCCCTTG 59
OY 2600 AGTTCCTGCGCAGAGCCCGCCCTGCAACAAGTCCCTGATGAAGCCACCGGCGCTGACC 2659
DB 60 AGTTCCTGCGCAGAGCCCGCCCTGCAACAAGTCCCTGATGAAGCCACCGGCGCTGACC 119
OY 2660 AGGAGTCAGAGATGACCTGAGCGGCTTGCAGACATCGCTGAGAGCGCAGGCCCGGCAAC 2719
```


PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001312.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225477P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226688P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-023497P.
PR 25-SEP-2000; 2000US-023498P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-023637P.
PR 29-SEP-2000; 2000US-023637P.

PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241222P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249246P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0255678P.

(HUMA-) HUMAN GENOME SCI INC.
XX
PA
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-465460/50.

This Page Blank (uspto)

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 22, 2006, 00:47:45 ; Search time 13987.1 Seconds
(Without alignments)
12533.446 Million cell updates/sec

Title: US-09-555-342B-1_COPY_49_3183

Perfect score: 3135
Sequence: 1 atggggagaatcagacagag.....acaaagagctctctgtgtat 3135

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 48236798 seqs, 2795965780 residues

Total number of hits satisfying chosen parameters: 96473596

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Lifting filter 100%

Database :

EST:*
1: gb_est1:*
2: gb_est3:*
3: gb_est4:*
4: gb_est5:*
5: gb_est6:*
6: gb_hic:*
7: gb_est2:*
8: gb_est7:*
9: gb_est8:*
10: gb_est9:*
11: gb_g881:*
12: gb_g882:*
13: gb_g883:*
14: gb_g884:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3133.4	99.9	4378	14 AY406167	AY406167 Homo sapi
2	3055	97.4	4775	14 CR859100	CR859100 Pongo pyg
3	2965.2	94.6	3138	14 AY406168	AY406168 Pan trogl
4	2425	77.4	3147	14 AY406169	AY406169 Mus muscu
5	1013.2	32.3	3743	6 AK171713	AK171713 Mus muscu
6	1007.2	32.1	3874	6 AK050860	AK050860 Mus muscu
7	957	30.5	968	4 BX364606	BX364606 Mus muscu
8	940.2	30.0	1000	4 BX354921	BX354921 BX354921
9	883.4	28.2	912	3 BQ881649	BQ881649 AGENCOURT
10	873.2	27.9	904	3 BU931960	BU931960 AGENCOURT
11	863.8	27.6	1026	2 BM551766	BM551766 AGENCOURT
12	852.4	27.2	1072	4 BX354920	BX354920 BX354920
13	842.4	26.9	3165	14 AY415515	AY415515 Homo sapi
14	839.8	26.8	959	3 BU539080	BU539080 AGENCOURT
15	834	26.6	857	4 BX350128	BX350128 BX350128
16	827	26.4	916	3 BU846495	BU846495 AGENCOURT
17	819.4	26.1	951	4 BX364605	BX364605 BX364605
18	817.2	26.1	966	3 BU146764	BU146764 AGENCOURT
19	816.6	26.0	889	3 BU174449	BU174449 AGENCOURT

20	812.6	25.9	926	3 BQ706499	BQ706499 AGENCOURT
21	802.2	25.6	3168	14 AY415517	AY415517 Mus muscu
22	801.2	25.6	806	10 DR763338	DR763338 HESG4_151
23	797.8	25.4	899	4 BX327170	BX327170 BX327170
24	795.8	25.4	800	9 CX781945	CX781945 HESG3_10
25	789.4	25.2	1354	10 DV784833	DV784833 Hw. liver
26	788.2	25.1	1043	3 BQ072025	BQ072025 AGENCOURT
27	784.8	25.0	1132	7 BE746268	BE746268 601579756
28	781.8	24.9	886	3 BQ672632	BQ672632 AGENCOURT
29	781	24.9	933	3 BQ950768	BQ950768 AGENCOURT
30	762	24.3	774	2 B1760610	B1760610 603044980
31	761.2	24.3	804	2 BG747792	BG747792 602705293
32	760.6	24.3	881	3 BU163030	BU163030 AGENCOURT
33	757.2	24.2	865	7 BE260677	BE260677 601154188
34	755.8	24.1	784	2 BG763918	BG763918 602736888
35	755.4	24.1	988	3 BQ674703	BQ674703 AGENCOURT
36	753	24.0	930	1 AL566821	AL566821 AL566821
37	751.8	24.0	952	3 BQ676994	BQ676994 AGENCOURT
38	748	23.9	944	3 BU173111	BU173111 AGENCOURT
39	745	23.8	861	2 BG764061	BG764061 602737071
40	730.4	23.3	745	8 CN281631	CN281631 170004710
41	726.2	23.2	901	2 BG475554	BG475554 602491544
42	725.6	23.1	3126	14 AY415516	AY415516 Pan trogl
43	719.2	22.9	942	3 BU179876	BU179876 AGENCOURT
44	716.4	22.9	741	8 CN291626	CN291626 170004245
45	716	22.8	736	2 BF793662	BF793662 602254351

ALIGNMENTS

RESULT 1
AY406167
LOCUS
DEFINITION
Homo sapiens FARP1 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION
AY406167.1 GI:39762141
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.
1 (bases 1 to 3138)
Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarival, A.,
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Shteyn, J.J.,
Adams, M.D. and Cargill, M.
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
Science 302 (5652), 1960-1963 (2003)

JOURNAL
PUBMED
14671302
2 (bases 1 to 3138)
Clark, A.G., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Shteyn, J.J.,
Adams, M.D. and Cargill, M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
This sequence was made by sequencing genomic exons and ordering
them based on alignment.
Location/Qualifiers
1..3138
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1..3138
/gene="FARP1"
/locus_tag="HCM2467"

ORIGIN

Query Match 99.9%; Score 3133.4; DB 14; Length 3138;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3134; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGAAATATGAGCAGAGGCCGACCCCGATCAACGACTGGGGCCCCGGAAATTTGG 60
DB 1 ATGGGAAATATGAGCAGAGGCCGACCCCGATCAACGACTGGGGCCCCGGAAATTTGG 60
QY 61 GGGATAGTACCTTGGAAAGTGGACAGAAAGCGCCCCCAACCTTCAGGAAATCTGTTG 120
DB 61 GGGATAGTACCTTGGAAAGTGGACAGAAAGCGCCCCCAACCTTCAGGAAATCTGTTG 120
QY 121 TCCATCAAAATCCAGATGCTGATGACACCCAGAGAGCATTTGAAAGTTCACAAAGAGCT 180
DB 121 TCCATCAAAATCCAGATGCTGATGACACCCAGAGAGCATTTGAAAGTTCACAAAGAGCT 180
QY 181 CCTGGAGAGGTGCTGCTGATGCAAGTTTGCACCACTCTGAGAGGTGACTAT 240
DB 181 CCTGGAGAGGTGCTGCTGATGCAAGTTTGCACCACTCTGAGAGGTGACTAT 240
QY 241 TTTGGCTGAGTTTCTGATCAAAAAGATCAGCGGTGAGTCTCTTAAACCC 300
DB 241 TTTGGCTGAGTTTCTGATCAAAAAGATCAGCGGTGAGTCTCTTAAACCC 300
QY 301 ATTGTAAACAGATTAGAAAGCCAAAGCAGTTGTTTAAAGTTTGTGAAATTTCTT 360
DB 301 ATTGTAAACAGATTAGAAAGCCAAAGCAGTTGTTTAAAGTTTGTGAAATTTCTT 360
QY 361 CGGCTGACCAACAACAATCTCAAGAAAGTCAACAAGTACTGTTCCGCTGCAAGTGG 420
DB 361 CGGCTGACCAACAACAATCTCAAGAAAGTCAACAAGTACTGTTCCGCTGCAAGTGG 420
QY 421 AAGCAGAGCTTGGCTCAAGGCAAGTGGATGATGACACAGCGAGCTCTTGATT 480
DB 421 AAGCAGAGCTTGGCTCAAGGCAAGTGGATGATGACACAGCGAGCTCTTGATT 480
QY 481 TCACACATTGTGCAATCTGAGATTGGGGAATTTGATGAAAGCTTGAACAGAGACTTA 540
DB 481 TCACACATTGTGCAATCTGAGATTGGGGAATTTGATGAAAGCTTGAACAGAGACTTA 540
QY 541 GCMAAAATTAATATCACTCTGACAAAGCAGCTAGAGACAAATTCGTGAAATTTTCA 600
DB 541 GCMAAAATTAATATCACTCTGACAAAGCAGCTAGAGACAAATTCGTGAAATTTTCA 600
QY 601 CATTAACCACTTGGACAAACACAGAGAGATCAGATTTCCAGCTCTAGAGATTGCCGT 660
DB 601 CATTAACCACTTGGACAAACACAGAGAGATCAGATTTCCAGCTCTAGAGATTGCCGT 660
QY 661 CGGCTAGAGATGTATGGAATCCGGTTGCAACCGGCCAAGGACAGGAGCAAGAGATC 720
DB 661 CGGCTAGAGATGTATGGAATCCGGTTGCAACCGGCCAAGGACAGGAGCAAGAGATC 720
QY 721 AATCTGGCCGTTGCCAACAGGGAATTTAGTGTTCAGGGTTTCACTAAGATCAATGCC 780
DB 721 AATCTGGCCGTTGCCAACAGGGAATTTAGTGTTCAGGGTTTCACTAAGATCAATGCC 780
QY 781 TTTCAACTGGGCGCAAGGTCGGAAGCTGAGCTTCAAGAGAAAGCCTTTCATCAAGCTC 840
DB 781 TTTCAACTGGGCGCAAGGTCGGAAGCTGAGCTTCAAGAGAAAGCCTTTCATCAAGCTC 840
QY 841 CGGCGAGATGCAATATGTGCTACAGAGATACCTTGAATTTCTGATGGCAGTCCGGAT 900
DB 841 CGGCGAGATGCAATATGTGCTACAGAGATACCTTGAATTTCTGATGGCAGTCCGGAT 900
QY 901 TTTCTGCAAGTCTTTCTGAAAAATCTGTGTGAACATATGCTTTCTTTAAGCTTTTGA 960
DB 901 TTTCTGCAAGTCTTTCTGAAAAATCTGTGTGAACATATGCTTTCTTTAAGCTTTTGA 960
QY 961 GAGCCCAAAACAAAGCCCAAGCCGATCTCTTTAGCCGGGGTCAATCATTTGGATTAGT 1020
DB 961 GAGCCCAAAACAAAGCCCAAGCCGATCTCTTTAGCCGGGGTCAATCATTTGGATTAGT 1020
QY 1021 GGTGGAAGTCAAGAGAGTTCTGACTATGTTAAAGAGAGACATTAAGAGGTGAC 1080

DB 1021 GGTGGAAGTCAAGAGAGTTCTGACTATGTTAAAGAGAGACATTAAGAGGTGAC 1080
QY 1081 TTTGAAAGAGACACAGAAAGTTCAATCTATCCGAGCCTTGCTTACAGGCTACAGAA 1140
DB 1081 TTTGAAAGAGACACAGAAAGTTCAATCTATCCGAGCCTTGCTTACAGGCTACAGAA 1140
QY 1141 CTGAATTCGGAAGTCTGAGAGCTCTCAGCAGAGCAGCCCTTACATTTGGAGAGGT 1200
DB 1141 CTGAATTCGGAAGTCTGAGAGCTCTCAGCAGAGCAGCCCTTACATTTGGAGAGGT 1200
QY 1201 GCCGAATCTCAGAGGGGCGAGAGCTGCGGCGAGAGAAAGAACCGAAGTTTCCCGGG 1260
DB 1201 GCCGAATCTCAGAGGGGCGAGAGCTGCGGCGAGAGAAAGAACCGAAGTTTCCCGGG 1260
QY 1261 GAGCCGGGGTGCACCCGAGCCCTGCGGAGAGAAAGCCCGGGGTAAACAAGCAGGG 1320
DB 1261 GAGCCGGGGTGCACCCGAGCCCTGCGGAGAGAAAGCCCGGGGTAAACAAGCAGGG 1320
QY 1321 GACGAGCCGCTCGGCGCCACAGAGAAAGAGAGGTGCTTAAAGATAGAACCCAG 1380
DB 1321 GACGAGCCGCTCGGCGCCACAGAGAAAGAGAGGTGCTTAAAGATAGAACCCAG 1380
QY 1381 CAGAGTAAACCTCAGCGCCCGCAGCCAGCAGAGCTCCCTGACTGGCAGTCTCACTT 1440
DB 1381 CAGAGTAAACCTCAGCGCCCGCAGCCAGCAGAGCTCCCTGACTGGCAGTCTCACTT 1440
QY 1441 TCCGAGCTGCTGTGTAACCTCGAGAGGGGGAGTGGCCCTGCAAGCTGCTTGTCCC 1500
DB 1441 TCCGAGCTGCTGTGTAACCTCGAGAGGGGGAGTGGCCCTGCAAGCTGCTTGTCCC 1500
QY 1501 AACCTGAGCCCGGACACCAAGAGGCTCTCCCTTGATCAGCCGCTGCTGAATACAG 1560
DB 1501 AACCTGAGCCCGGACACCAAGAGGCTCTCCCTTGATCAGCCGCTGCTGAATACAG 1560
QY 1561 GCCTGCCCCCGGACGGAATGAGATGAGGGCCGAGAGAAAGATTTCCCACTGATAA 1620
DB 1561 GCCTGCCCCCGGACGGAATGAGATGAGGGCCGAGAGAAAGATTTCCCACTGATAA 1620
QY 1621 GGGTACTTATATGCTTAAGGAAGTGTCTACACAGAGCCAAATATGTAAGATCTCGAA 1680
DB 1621 GGGTACTTATATGCTTAAGGAAGTGTCTACACAGAGCCAAATATGTAAGATCTCGAA 1680
QY 1681 GTTATCACTTCGTGGTTTCAAGACAGTGAACAAAGAGAGCCATGCGGAAAGACTGG 1740
DB 1681 GTTATCACTTCGTGGTTTCAAGACAGTGAACAAAGAGAGCCATGCGGAAAGACTGG 1740
QY 1741 AAAAGTCTATATTCGCCGAATTTTGAACCTTTGCAAAATTTCTAATTTTCTCAAG 1800
DB 1741 AAAAGTCTATATTCGCCGAATTTTGAACCTTTGCAAAATTTCTAATTTTCTCAAG 1800
QY 1801 GAAATTGAGCAAGACTTGGCCCTGAGGGAAGGCGCTCAATAGCCCAATTCAGATTAC 1860
DB 1801 GAAATTGAGCAAGACTTGGCCCTGAGGGAAGGCGCTCAATAGCCCAATTCAGATTAC 1860
QY 1861 CAAAGATCGGAGATGTCACTGTAAGAACATTCAGGGCATGAAGCACTGGCGGCTCAC 1920
DB 1861 CAAAGATCGGAGATGTCACTGTAAGAACATTCAGGGCATGAAGCACTGGCGGCTCAC 1920
QY 1921 CTGTGAGACACAGAGGCTTTGAGAGCTCTGAGAAATGGAATCAAGAGCTCCGGCGG 1980
DB 1921 CTGTGAGACACAGAGGCTTTGAGAGCTCTGAGAAATGGAATCAAGAGCTCCGGCGG 1980
QY 1981 CTGGAAGACTTCTGAGAGACTTTGAGCTGCAAGAGGTGTGTAACCTACCGCTCAACAC 2040
DB 1981 CTGGAAGACTTCTGAGAGACTTTGAGCTGCAAGAGGTGTGTAACCTACCGCTCAACAC 2040
QY 2041 TTTCTCTGCGGCACTGCAAGCGGCTCATGACATCAACAGAGTCTTGAAGCGGCTGTGC 2100
DB 2041 TTTCTCTGCGGCACTGCAAGCGGCTCATGACATCAACAGAGTCTTGAAGCGGCTGTGC 2100
QY 2101 AAACACACCCGCGGAGCCGACCTTCAAGGACTGCGAGCGCTTTGGCAGAGATC 2160

Db	2101	AAACAACACCCGCCAGCCAGCCCACTTTCAGAGACTCCGACCGCCCTTTGGCAGAGATC	2160
Qy	2161	ACGAGATGTTGGCAACAGCTCCACGGTACGATGATCAAGATGAGAAATTTCCAGAGCTG	2220
Db	2161	ACGGAGATGTTGGCAACAGCTCCACGGTACGATGATCAAGATGAGAAATTTCCAGAGAGCTG	2220
Qy	2221	CACGAACTCAAGAAAGATTTGATTGGCATTTGCAATCTTTGGTTCGGGAAAGGAGTTTC	2280
Db	2221	CACGAACTCAAGAAAGATTTGATTGGCATTTGCAATCTTTGGTTCGGGAAAGGAGTTTC	2280
Qy	2281	ATCCGCTCTGGGCAAGCTCTGAGCAAGCTCTCGGGGAAAGGGGCTCCAGCAGCGATGTTCTTC	2340
Db	2281	ATCCGCTCTGGGCAAGCTCTGAGCAAGCTCTCGGGGAAAGGGGCTCCAGCAGCGATGTTCTTC	2340
Qy	2341	CTGTTTCAACGAGCTCTGATATACAGAGCCGGGGGCTGACGGCTCCAAATCAAGTTTAA	2400
Db	2341	CTGTTTCAACGAGCTCTGATATACAGAGCCGGGGGCTGACGGCTCCAAATCAAGTTTAA	2400
Qy	2401	GTCCACGGGGCAGCTCCGGCTCTATGGCATGACGATTTGAGAGAGCGAAAGCAGATGGGGG	2460
Db	2401	GTCCACGGGGCAGCTCCGGCTCTATGGCATGACGATTTGAGAGAGCGAAAGCAGATGGGGG	2460
Qy	2461	GTGCCCCCACTGCTGACCTCCGGGGGCGAGGGGCGAGTCAATCATCTGTTGGCCGCACTTCT	2520
Db	2461	GTGCCCCCACTGCTGACCTCCGGGGGCGAGGGGCGAGTCAATCATCTGTTGGCCGCACTTCT	2520
Qy	2521	CGGTCGAGATGAGAGATGGGTTGAGGACATCCAGATGGCCATTGAACCTGGCGAGAG	2580
Db	2521	CGGTCGAGATGAGAGATGGGTTGAGGACATCCAGATGGCCATTGAACCTGGCGAGAG	2580
Qy	2581	AGCAGCAGCCCCGCCCTGAGTTCTGGGCCAGCAGCCCCCTGACAAACAATGCCCTGAT	2640
Db	2581	AGCAGCAGCCCCGCCCTGAGTTCTGGGCCAGCAGCCCCCTGACAAACAATGCCCTGAT	2640
Qy	2641	GAAGCCAACGGGGGCTGACAGAGATGACAGATGACCTGACCGCTCCGGGCAATCGCTG	2700
Db	2641	GAAGCCAACGGGGGCTGACAGAGATGACAGATGACCTGACCGCTCCGGGCAATCGCTG	2700
Qy	2701	GAGCGCCAGGCCCCCGCACCGGGCAACAATGTTGTCAGTGTGTGGCAACCGCAACCC	2760
Db	2701	GAGCGCCAGGCCCCCGCACCGGGCAACAATGTTGTCAGTGTGTGGCAACCGCAACCC	2760
Qy	2761	AGCGTCTCCATGTTGGAATTCAGCATGCGAGTGAAGAAATCAGTTGTCTGAAACCTGCTG	2820
Db	2761	AGCGTCTCCATGTTGGAATTCAGCATGCGAGTGAAGAAATCAGTTGTCTGAAACCTGCTG	2820
Qy	2821	AGGAAATTCAAAAACAGAACGGGTGGGCAAGAGCTGTGGGTGTTTCAAAATCTTCTGC	2880
Db	2821	AGGAAATTCAAAAACAGAACGGGTGGGCAAGAGCTGTGGGTGTTTCAAAATCTTCTGC	2880
Qy	2881	CTGTTTCTTCAAAATCAACACGAGCAATCATCCCTTGGCAGGCTGCTCTGCTGGGC	2940
Db	2881	CTGTTTCTTCAAAATCAACACGAGCAATCATCCCTTGGCAGGCTGCTCTGCTGGGC	2940
Qy	2941	TACTGCTCACCATCCCTCTGAGTCCGAGAACATCCAGAAAGACTAGTGTCAAGCTG	3000
Db	2941	TACTGCTCACCATCCCTCTGAGTCCGAGAACATCCAGAAAGACTAGTGTCAAGCTG	3000
Qy	3001	CACTTCAAGTCCCAAGTCTACTACTTTCAGGGCGGAAAGCGATACGTTTGAAAGGTGG	3060
Db	3001	CACTTCAAGTCCCAAGTCTACTACTTTCAGGGCGGAAAGCGATACGTTTGAAAGGTGG	3060
Qy	3061	ATGGAAGTGAATCCGAGTGCACCAAGCTCTGCTGGCAACCCCAAGTGTGATGCATAA	3120
Db	3061	ATGGAAGTGAATCCGAGTGCACCAAGCTCTGCTGGCAACCCCAAGTGTGATGCATAA	3120
Qy	3121	GAGTCTCTTGTGTAT	3135
Db	3121	GAGTCTCTTGTGTAT	3135

LOCUS	CR859100	4775 bp	mRNA	linear	HTC 12-NOV-2004
DEFINITION	Pongo pygmaeus mRNA; CDNA DKFZp469C092 (from clone DKFZp469C092).				
ACCESSION	CR859100				
VERSION	CR859100.1	GI:55729110			
KEYWORDS	HTC.				
SOURCE	Pongo pygmaeus (orangutan)				
ORGANISM	Pongo pygmaeus				
AUTHORS	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiinae; Pongo.				
CONSRMT	1 (bases 1 to 4775)				
TITLE	Bloecher H., Boecher M., Brandt P., Mewes H.W., Weill B., Amid C., Oeangger A., Fobio G., Han M. and Wiemann S.				
JOURNAL	The German CDNA Consortium				
COMMENT	Direct Submission Submitted (12-NOV-2004) MIPS, Ingolstaedter Landstr.1, D-85764 Neuherberg, GERMANY Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; Sequenced by GBR (National Research Centre for Biotechnology Ltd., Braunschweig/Germany) within the CDNA sequencing consortium of the German Genome Project. This clone (DKFZp469C092) is available at the RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany. Please contact RZPD for ordering: http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneid=DKFZp469C092 Further information about the clone and the sequencing project is available at http://mips.gsf.de/projects/cdna/.				
FEATURES	Location/Qualifiers				
SOURCE	1..4775				
gene	/organism="Pongo pygmaeus"				
CDS	/mol_type="mRNA"				
	/db_xref="taxon:9600"				
	/clone="DKFZp469C092"				
	/tissue_type="kidney"				
	/clone_id="469 (synonym: pkiol). Vector pSPORT1_Sfi; host: DH10B; sites SfiIa & SfiIb"				
	/dev_stage="adult"				
	/note="FERM, RhogEF, and plectroin domain protein 1 (Homo sapiens)"				
	1..4775				
	/gene="DKFZp469C092"				
	100..3237				
	/gene="DKFZp469C092"				
	/product="hypothetical protein"				
	/codon_start=1				
	/protein_id="CAH91292.1"				
	/db_xref="GI:55729111"				
	/translation="MGEIEQPTPGSLGAPENSGISTLERGQKPEPTPSGKLVISIKI QMDDIDGAEVEFQAPAKVILDAVCHNLIVSGDYGELCDPHKKTITWLDILKPLV KQIRRPDEAVFVKKFPFDPDLOELFRVYFALOVKDCGRITCDWDTALALI SHVQSIGDPDADREHAKKKYIPQDADLVLEFHHNIGCPASDPGLIIT ARLEMTGIRLHPAKDBEGTKINLVANNTIIVPGSTKINNAWAVRGLSPFRKRF LILKRPANSAVQDTLEFLWASBDFPCKSEPKIVENHAFRLREEPKPKPVLFSGG ASIFRGRTQKQVLDVYEGEGHKKVQFERGSHSIHSIRLSAQPTIELYSVLREQSOOS ASIFLFGCAISPGSQCOQCEKPKVSGEESHSPVPRSPAGNKQADGASAPTEB EEEVVKQRTQSOOKPPOPPDSTGSLTSGPHLSVNSQGVAVANVTLSNLSPTQK ASPLIFPLINDQACPTDDBDGRKRRKPRDXAVFLAKVSTERTYLIKDLVITSMF OSASVEDAMPALAKSLIFNPFERLKFHTNPLAKEIQRALMBEGSNADIRYQRIG DVMKIKQMKHLAVHLWKSHSEALENLIKSSRLLENCRDPELOKVCYLPLNTLL LRLPRHMYKQVLERLCKHPSSHADFPDRCALAGITEMVAQHQTMKMEFQKL HEKKKLDIGIDNLVVPGRERILRGLSLKSGKALQORMFQNDVLVLYTRGLTASQD FKHGQGLPYGMITIKESBDEWGHCTLGTGQOSIIVAASSSEMEKVEDIOMADID LAKNSLDAPEPLASSPDKNSPDENALAOESDLSASRTLERQAPRGMTHYV CRRNNTSVKAVDVSVAVENQSLNLRKFNKSGKQCLVYFNFCLPFKSHQDNHP LALPLIGYSLTIPTSESNTHKDVVFELFKSHVYVFRASEYTFPMWEIVRSATPS ASRVHVSXSHKESLVY"				
ORIGIN	Query Match	97.4%	Score 3055;	DB 6;	Length 4775;
	Best Local Similarity	98.4%	Pred. No. 0;		
	Matches 3085;	Conservative	0;	Mismatches	50;
				Indels	0;
				Gaps	0;

QY 1 ATGGAGAAATAGAGACAGAGCCGACCCAGATCAAGACTGGGGGCCCCGAAATTCG 60
Db 100 ATGGAGAAATAGAGAGAGCCGACCCAGATCAAGACTGGGGGCCCCGAAATTCG 159
QY 61 GGGATCAGTACCTTGGACCTGGACAGAGCCGCCCAACACTTCAGAGAAATCTCGTG 120
Db 160 GGGATCAGTACCTTGGACCTGGACAGAGCCGCCCAACACTTCAGAGAAATCTCGTG 219
QY 121 TCCATCAAAATCCAGATGCTGGATGACACCCAGAGAGCTTTGAAGTTCCAGAAAGCT 180
Db 220 TCCATCAAAATCCAGATGCTGGATGACACCCAGAGAGCTTTGAAGTTCCAGAAAGCT 279
QY 181 CCTGGAGAGGTGCTGCTGATGACAGTTTGCACCACTCACTCGTGAAGGTGACTAT 240
Db 280 CCGGGAGAGGTGCTGCTGATGACAGTTTGCACCACTCACTCGTGAAGGTGACTAT 339
QY 241 TTTGGCTCGAGTTTCTCTGATCAAAAAGATCACGGTGGCTGGATCTCTTAAACC 300
Db 340 TTTGGCTCGAGTTTCTCTGATCAAAAAGATCACGGTGGCTGGATCTCTTAAACC 399
QY 301 ATTGTGAACAGATTGAAGGCCAAAGCAGTTGTTTAAGTTTGTGGAATTCCTT 360
Db 400 CTGTGGAACAGATCAGAGGCCAAAGCAGTTGTTTGAAGTTTGTGGAATTCCTT 459
QY 361 CCGCTGACCAACAACACTCCAAAGAACTCAAGGTAACCTGTTCGCGCTGCAGGTG 420
Db 460 CCGCTGACCAACAACACTCCAAAGAACTCAAGGTAACCTGTTCGCGCTGCAGGTG 519
QY 421 AAGCAGACTTGGCTCAAGCAGGTGACGTGTAATGACACAGCCGACGCTCTTGAT 480
Db 520 AAGCAGACTTGGCTCAAGCAGGTGACGTGTAATGACACAGCCGACGCTCTTGAT 579
QY 481 TCACACATTTGCAATCTGATTTGGGGATTTTGTAGAAAGCTTTGACAGAGACCTTA 540
Db 580 TCACACATTTGCAATCTGATTTGGGGATTTTGTAGAAAGCTTTGACAGAGACCTTA 639
QY 541 GCAAAAATTAATACATCTCTGACAGACGCACTAGAGACAAATTCGTGAATTCAC 600
Db 640 GCAAAAATTAATACATCTCTGACAGACGCACTAGAGACAAATTCGTGAATTCAC 699
QY 601 CATTAACAATTGGACAAACACAGCAGATTCAGATTTCCAGCTCTTGAAGATTCGCC 660
Db 700 CATTAACAATTGGACAAACACAGCAGATTCAGATTTCCAGCTCTTGAAGATTCGCC 759
QY 661 CCGCTAGAGATGATGGAATCCGGTTGCAACCCGACCAAGACAGAGAGGCAAGAGATC 720
Db 760 CCGCTAGAGATGATGGAATCCGGTTGCAACCCGACCAAGAGAGGCAAGAGATC 819
QY 721 AATCTGACCTTGCACACAGGAAATCTAGTGTTCAGGGTTTCACTAAGATCAATGCC 780
Db 820 AATCTGACCTTGCACACAGGAAATCTAGTGTTCAGGGTTTCACTAAGATCAATGCC 879
QY 781 TTCAATCGGGCCAAAGTGGCGGAAGCTGAGCTTCAAGAGAAAGGCTTTCTCATCAAGCTC 840
Db 880 TTCAATCGGGCCAAAGTGGCGGAAGCTGAGCTTCAAGAGAAAGGCTTTCTCATCAAGCTC 939
QY 841 CCGCAGATGCAATAGTGTGCTACAGAGATTCCTGAATTCCTGATGCGCATCGGAT 900
Db 940 CCGCAGATGCAATAGTGTGCTACAGAGATTCCTGAATTCCTGATGCGCATCGGAT 999
QY 901 TTCTGCAAGTCTTCTGGAATCTGTGTGAACATCAGCTCTTCTTGAATCTTTTGA 960
Db 1000 TTCTGCAAGTCTTCTGGAATCTGTGTGAACATCAGCTCTTCTTGAATCTTTTGA 1059
QY 961 GAGCCCAAAACCAAGCCCAAGCCCTCTCTTAAAGCGGGGTCATCTTCCGTTCAAG 1020
Db 1060 GAGCCCAAAACCAAGCCCAAGCCCTCTCTTAAAGCGGGGTCATCTTCCGTTCAAG 1119
QY 1021 GGTGCGACTCAGAGAGGTTCTGACTATGTTAAAGAAAGAGACATTAAGAGGTGAC 1080
Db 1120 GGTGCGACTCAGAGAGGTTCTGACTATGTTAAAGAAAGAGACATTAAGAGGTGAC 1179
QY 1081 TTGAAAGAGACAGACAGATTCATTTATCCGAGCCTTGTCTTCAACCTTACAGAA 1140

Db 1180 TTGAAAGAGACAGACAGATTCATTTATCCGAGCCTTGTCTTCAACCTTACAGAA 1239
QY 1141 CTGAATTCGGAATGCTGAGAGAGTCTACAGAGACCAAGCTTCAATTTGAGAGAGGT 1200
Db 1240 CTGTATTCGGAAGTACTGAGAGAGTCTACAGAGAGCCAGCTTCAATTTGAGAGAGGT 1299
QY 1201 GCCGAATCTCAGAGGGGCGAGAGCTGCGCGAGAGAAAGAAACCGAAGGTTTCCGCGGG 1260
Db 1300 GCCGAATCTCAGAGGGGCGAGAGCTGCGCGAGAGAAAGAAACCGAAGGTTTCCGCGGG 1359
QY 1261 GAGCGGGGTGCAACCCGAGCCTGCGCGAGAGAGAGCCCGCGGTAAACAAGCGCG 1320
Db 1360 GAGCGGGGTGCAACCCGAGCCTGCGCGAGAGAGAGCCCGCGGTAAACAAGCGCG 1419
QY 1321 GACGAGCGGCTCGCGCGCGCAAGAGAGAGAGAGAGTGTGTAAGATGGAACCGAG 1380
Db 1420 GACGAGCGGCTCGCGCGCGCAAGAGAGAGAGAGAGTGTGTAAGATGGAACCGAG 1479
QY 1381 CAGAGTAAACTCAGCGCCCGAGCCAGCAAGCAAGGCTCCCTGACTGGCAGTCTCACTT 1440
Db 1480 CAGAGTAAACTCAGCGCCCGAGCCAGCAAGCAAGGCTCCCTGACTGGCAGGCTTCACTT 1539
QY 1441 TCGAGCTGTCTGTGAATCTGCAAGGGGGAGTGGCTCTGCAACCTGTCTTCTCC 1500
Db 1540 TCGAGCTGTCTGTGAATCTGCAAGGGGGAGTGGCTCTGCAACCTGTCTTCTCC 1599
QY 1501 AACCTGAGCCCGAGCAACAGAGAGGCTCTCCCTTGAATGAGCCGCTGATGACAG 1560
Db 1600 AACCTGAGCCCGAGCAACAGAGAGGCTCTCCCTTGAATGAGCCGCTGATGACAG 1659
QY 1561 GCTGCCCCCGAGCAAGAGAGGCTCTCCCTTGAATGAGCCGCTGATGACAG 1620
Db 1660 GCTGCCCCCGAGCAAGAGAGGCTCTCCCTTGAATGAGCCGCTGATGACAG 1719
QY 1621 GCGTACTTATGCTAAGAGAGTGTCTTACACCGAGCGAACTATTTGAAGATCTCGAA 1680
Db 1720 GCGTACTTATGCTAAGAGAGTGTCTTACACCGAGCGAACTATTTGAAGATCTCGAA 1779
QY 1681 GTTATCACTTGGGGTTTGAAGCAAGTGAACAAAGAGAGAGCCATGCGGAGCACTG 1740
Db 1780 GTTATCACTTGGGGTTTGAAGCAAGTGAACAAAGAGAGAGCCATGCGGAGCACTG 1839
QY 1741 AAAAGTCTATATTCGCCAATTTTGAACCTTTGCAAAATTTCAATTAATTTCTCAAG 1800
Db 1840 AAAAGTCTATATTCGCCAATTTTGAACCTTTGCAAAATTTCAATTAATTTCTCAAG 1899
QY 1801 GAAATGAGCAAGACTTCCCTGTGGGAAAGCCGCTCAAAATGCCCAATTCAGAGTTAC 1860
Db 1900 GAAATGAGCAAGACTTCCCTGTGGGAAAGCCGCTCAAAATGCCCAATTCAGAGTTAC 1959
QY 1861 CAAAGATCGGGAGTCACTGCTGAAGAACTTCAAGGGGATGAGACCTGGCGGCTCAC 1920
Db 1960 CAAAGATCGGGAGTCACTGCTGAAGAACTTCAAGGGGATGAGACCTGGCGGCTCAC 2019
QY 1921 CTGTGAGACCAAGCAGGCTTGGAGGCTCTGAGAAATGGAATCAAGACTTCCGCGG 1980
Db 2020 CTGTGAGACCAAGCAGGCTTGGAGGCTCTGAGAAATGGAATCAAGACTTCCGCGG 2079
QY 1981 CTGAGAACTTCTGCAAGAGACTTTGAGCTGCAAGAGGTGTGTTACTTACCGCTCAAC 2040
Db 2080 CTGAGAACTTCTGCAAGAGACTTTGAGCTGCAAGAGGTGTGTTACTTACCGCTCAAC 2139
QY 2041 TTCTCTCGCGGCACTGCAAGGCTCATGCACTCAAGCAGGTCCTGAGCGGCTGTG 2100
Db 2140 TTCTCTCTCGCGGCACTGCAAGGCTCATGCACTCAAGCAGGTCCTGAGCGGCTGTG 2199
QY 2101 AAACACCAACCGCGAGCCAGCCGACTTCAAGGACTGCGAGCCGCTTTGGCAGAGATC 2160
Db 2200 AAACACCAACCGCGAGCCAGCCGACTTCAAGGACTGCGAGCCGCTTTGGCAGAGATC 2259
QY 2161 ACGAGATGTGCAAGCTTCAAGGATGATCAAGATGAGAAATTTCCAGAGCTG 2220

```

Db 2260 ACGAGATGTTGTCAGACAGCTCCAGGTACGATGATCAAGATGAGAAATTTCCAGAGCTG 2319
Qy 2221 CAGAACTTCAAGAAAGATTTGATTTGGCATTTGACATCTTGTGTTCGGGAAAGGAGTTC 2280
Db 2320 CACGAACTCAAGAAAGATTTGATTTGGCATTTGACATCTTGTGTTCGGGAAAGGAGTTC 2319
Qy 2281 ATCCGTCTGGGACAGCTCAGCAAGCTCTCGGGAAAGGGCTCCAGCAGCGCATTTCTTC 2340
Db 2280 ATCCGTCTGGGACAGCTCAGCAAGCTCTCGGGAAAGGGCTCCAGCAGCGCATTTCTTC 2439
Qy 2341 CTGTTCACGACGCTCTGTATACAGAGCCGGGGGCTGACGGGCTCCATCACTTTAAA 2400
Db 2440 CTGTTCACGACGCTCTGTATACAGAGCCGGGGGCTGACGGGCTCCATCACTTTAAA 2499
Qy 2401 GTCCACGGGACAGCTCCGGCTCTATAGCATGACATTTAGAGAGACGAAAGAGTGGGGG 2460
Db 2500 GTCCACGGGACAGCTCCGGCTCTATAGCATGACATTTAGAGAGACGAAAGAGTGGGGG 2559
Qy 2461 GTGCCCCAGTGCCTGACCTCCGGGGCAGCGGAGTCCATCATGTGGCCGAGTTCT 2520
Db 2560 GTGCCCCAGTGCCTGACCTCCGGGGCAGCGGAGTCCATCATGTGGCCGAGTTCT 2619
Qy 2521 CGGTCCGAGATGAGAAAGTGGTTAGAGACATCCAGATGGCCATTTGACCTGGCGAGAG 2580
Db 2620 CGGTCTGAGATGAGAAAGTGGTTAGAGACATCCAGATGGCCATTTGACCTGGCGAGAG 2679
Qy 2581 AGCAGACGCCCCCGCTGAGTTCCTGGCCAGCGCCCCCTGACAAAGTCCCTGAT 2640
Db 2680 AACAGACGCTCCGCTGAGTTCCTGGCCAGCGCCCCCTGACAAAGTCCCTGAT 2739
Qy 2641 GAAGCCACCGCGCTGACAGAGATGAGAGATGAGAGTCCGCGCATTCCTGAT 2700
Db 2740 GAAGCCACCGCGCTGACAGAGATGAGAGATGAGAGTCCGCGCATTCCTGAT 2799
Qy 2701 GAGCGCCAGGCGCCGCGACCGCGGACAAATGCTGTGCTGGCAACCCGACACC 2760
Db 2800 GAGCGCCAGGCGCCGCGACCGCGGACAAATGCTGTGCTGGCAACCCGACACC 2859
Qy 2761 AGGCTCTCCATGTTGAGCTTCAAGATGCGAGTGGAGAAATCAGTTGTCTGAAAACCTGCTG 2820
Db 2860 AGGCTCTCCATGTTGAGCTTCAAGATGCGAGTGGAGAAATCAGTTGTCTGAAAACCTGCTG 2919
Qy 2821 AGGAAATTTCAAAAACAGCAACGAGTGGAGAGTGTGGTGTTCACAAACTTCTGC 2880
Db 2920 AGGAAATTTCAAAAACAGCAACGAGTGGAGAGTGTGGTGTTCACAAACTTCTGC 2979
Qy 2881 CTGTCTCTTACAAATACACACAGAACATATCCCTTGCAGCTGCTGTGCTGGC 2940
Db 2980 CTGTCTCTTACAAATACACACAGAACATATCCCTTGCAGCTGCTGTGCTGGC 3039
Qy 2941 TACTCGCTCACTCCCTCTGAGTCCGAGAACATCCAGAAAGATGAGTGTTCAGCTG 3000
Db 3040 TACTCGCTCACTCCCTCTGAGTCCGAGAACATCCAGAAAGATGAGTGTTCAGCTG 3099
Qy 3001 CACTTCAAGTCCCAAGCTCTACTCTTCAAGGCGGAAAGAGATCAAGTTCGAAGGTG 3060
Db 3100 CACTTCAAGTCCCAAGCTCTACTCTTCAAGGCGGAAAGAGATCAAGTTCGAAGGTG 3159
Qy 3061 ATGAAAGTATCCGACAGTGCACCAAGCTCTGCTCGACCCCACTGTGAGCCAAA 3120
Db 3160 ATGAAAGTATCCGACAGTGCACCAAGCTCTGCTCGACCCCACTGTGAGCCAAA 3219
Qy 3121 GAGTCTCTTGTGAT 3135
Db 3220 GAGTCTCTTGTGAT 3234

```

```

RESULT 3
AY406168 3138 bp DNA linear GSS 15-DEC-2003
LOCUS Pan troglodytes PAR1 gene, VIRUAL TRANSCRIPT, partial sequence,
DEFINITION genomic survey sequence.
ACCESSION AY406168

```

```

VERSION AY406168.1 GI:39762142
KEYWORDS GSS
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Pan.
REFERENCE 1 (bases 1 to 3138)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,
Todd,M.A., Tanenbaum,D.M., Civeillo,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Smitsky,J.J.,
Adams,M.D. and Cargill,M.
Inferred nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 3138)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,
Todd,M.A., Tanenbaum,D.M., Civeillo,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Smitsky,J.J.,
Adams,M.D. and Cargill,M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
This sequence was made by sequencing genomic exons and ordering
them based on alignment
COMMENT
FEATURES
Location/Qualifiers
source 1..3138
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
<1..>3138
/feature="PAR1"
/locus_tag="HCM2467"
ORIGIN
Query Match 94.6%; Score 2965.2; DB 14; Length 3138;
Best Local Similarity 94.8%; Pred. No. 0;
Matches 2973; Conservative 0; Mismatches 162; Indels 0; Gaps 0;
Qy 1 ATGGAGAAATAGAGCAGAGGCGGACCCAGATCAAGTGGGGGCGGAAATTCG 60
Db 1 ATGGAGAAATAGAGCAGAGGCGGACCCAGATTCGAGTGGGGGCGGAAATTCG 60
Qy 61 GGGATCAGTACCTTGGAACTGAGCAGAGCCGCCCAACCTTCAGAGAAATCTGTG 120
Db 61 GGGATCAGTACCTTGGAACTGAGCAGAGCCGCCCAACCTTCAGAGAAATCTGTG 120
Qy 121 TCCATCAAAATTCAGATGCTGATGACACCCAGAGGCAATTTGAAGTTCACAAAGAGCT 180
Db 121 TCCATCAAAATTCAGATGCTGATGACACCCAGAGGCAATTTGAAGTTCACAAAGAGCT 180
Qy 181 CTGGAGAGGTGCTGTGATGATGACAGTTTGCAACACTCAACTCTGTGAAGGTGACTAT 240
Db 181 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 240
Qy 241 TTTGGCTCGAGTTTCTCTGATCAAAAGATACAGGTGTGCTGATCTCTTAAACC 300
Db 241 TTTGGCTCGAGTTTCTCTGATCAAAAGATACAGGTGTGCTGATCTCTTAAACC 300
Qy 241 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 300
Db 241 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 300
Qy 301 ATTGTGAACAGATTGAAGGCCAAAGCACTGTTGTTAAGTTTGTGGTGAATCTTT 360
Db 301 ATTGTGAACAGATTGAAGGCCAAAGCACTGTTGTTAAGTTTGTGGTGAATCTTT 360
Qy 361 CCGCTGACCAACACAACTCCAAAGAAAGTCAAGGTAAGTCTGTGCGCTGACAGTG 420
Db 361 CCGCTGACCAACACAACTCCAAAGAAAGTCAAGGTAAGTCTGTGCGCTGACAGTG 420
Qy 421 AACGAGACTTGGCTCAAGCAGGTTGAGTGAATGACACCGAGGCACTCTTGATT 480
Db 421 AACGAGACTTGGCTCAAGCAGGTTGAGTGAATGACACCGAGGCACTCTTGATT 480
Qy 481 TCACACATGTGATCTGATTTGGGATTTTGTATGAAGCTTGAAGAGACACTTA 540

```

```
Db      ||||| 481 TCACATTTGTGCAATCTGAGATTGGGGATTTTGAAGACTTTGGACAGAGGACCTTA 540
Qy      ||||| 541 GCAAAAAATTAATATCTACTCTGAGCAAGAGCACTAGAGGACAAATCGTGAAATTTCC 600
Db      ||||| 541 GCAAAAAATTAATATCTACTCTGAGCAAGAGCACTAGAGGACAAATCGTGAAATTTCC 600
Qy      ||||| 601 CATTAACAATTTGAGCAAAACACAGCAGAGATCAGATTTCCAGCTCTAGAGATTCGCCG 660
Db      ||||| 601 CATTAACAATTTGAGCAAAACACAGCAGAGATCAGATTTCCAGCTCTAGAGATTCGCCG 660
Qy      ||||| 661 CGGCTAGAGATGATGGAATCCGGTTGCAACCGGCCAAGGACAGGAGGCAAGGAATC 720
Db      ||||| 661 CGGCTAGAGATGATGGAATCCGGTTGCAACCGGCCAAGGACAGGAGGCAAGGAATC 720
Qy      ||||| 721 AATCTGGCCGTTGCCCAACAGGGAATTTCTAGTTTCAAGGTTTCACTAAGATCAATGCC 780
Db      ||||| 721 AATCTGGCCGTTGCCCAACAGGGAATTTCTAGTTTCAAGGTTTCACTAAGATCAATGCC 780
Qy      ||||| 781 TTCAACTGGGCGCAAGGTGCGGAGCTGAGCTTCAAGAGAAAGCGCTTTCATCAAGCTC 840
Db      ||||| 781 TTCAACTGGGCGCAAGGTGCGGAGCTGAGCTTCAAGAGAAAGCGCTTTCATCAAGCTC 840
Qy      ||||| 841 CGGCAAGATGCAATATGTGCGTAACAAGATACCTTGAAATTCCTGATGCGCAATGCGAT 900
Db      ||||| 841 CGGCAAGATGCAATATGTGCGTAACAAGATACCTTGAAATTCCTGATGCGCAATGCGAT 900
Qy      ||||| 901 TTCTGCAAGCTCTTCTGAGAAAATCTGTGTGAACATCAATGCTTCTTGAATCTTTTGA 960
Db      ||||| 901 TTCTGCAAGCTCTTCTGAGAAAATCTGTGTGAACATCAATGCTTCTTGAATCTTTTGA 960
Qy      ||||| 961 GAGCCCAACCAAGGCCCAAGCCGTCCTTTAGCGGGGTCATCATTTCCGTTCACT 1020
Db      ||||| 961 GAGCCCAACCAAGGCCCAAGCCGTCCTTTAGCGGGGTCATCATTTCCGTTCACT 1020
Qy      ||||| 1021 GGTGGAATCTCAGAAAGGTTCTGACTATGTTAAAGAGAGACATAAGAGGTGACG 1080
Db      ||||| 1021 GGTGGAATCTCAGAAAGGTTCTGACTATGTTAAAGAGAGACATAAGAGGTGACG 1080
Qy      ||||| 1081 TTTGAAAGAGAGCAACAGCAAGATTTCAATTCGAGACCTTGCTTCAAGCCTTACAGA 1140
Db      ||||| 1081 TTTGAAAGAGAGCAACAGCAAGATTTCAATTCGAGACCTTGCTTCAAGCCTTACAGA 1140
Qy      ||||| 1141 CTGAATTCGGAATGTCTGAGCAGCTCTCAGCAGAGCAACAGCCTTACATTTGAGAGAGT 1200
Db      ||||| 1141 CTGAATTCGGAATGTCTGAGCAGCTCTCAGCAGAGCAACAGCCTTACATTTGAGAGAGT 1200
Qy      ||||| 1201 GCCGAATCTTCAGGGGGCCAGAGCTGCGGCGAGAAAGAACCGAAGGTTTCCGCCGG 1260
Db      ||||| 1201 GCCGAATCTTCAGGGGGCCAGAGCTGCGGCGAGAAAGAACCGAAGGTTTCCGCCGG 1260
Qy      ||||| 1261 GAGCCGGGGGTGCAACCGAGCCTTGCGCGAGAGAAAGCCCGGGGTAAACAAGCGGG 1320
Db      ||||| 1261 GAGCCGGGGGTGCAACCGAGCCTTGCGCGAGAGAAAGCCCGGGGTAAACAAGCGGG 1320
Qy      ||||| 1321 GACGAGCCGCTCTGCGGCCCAAGAGAGAGAGAGAGTCTGTTAAGATAGAACCCAG 1380
Db      ||||| 1321 GACGAGCCGCTCTGCGGCCCAAGAGAGAGAGAGAGTCTGTTAAGATAGAACCCAG 1380
Qy      ||||| 1381 CAGAGTAACTTCAAGCCCCCGCAGCCAGCAAGAGCTCCTGACTGCACTGCTCACTT 1440
Db      ||||| 1381 CAGAGTAACTTCAAGCCCCCGCAGCCAGCAAGAGCTCCTGACTGCACTGCTCACTT 1440
Qy      ||||| 1441 TCCGACTGTCTGTGAATCTGCAAGGGGGAGTGGCCCTGSCCAACGTGACTTGTCTCC 1500
Db      ||||| 1441 TCCGACTGTCTGTGAATCTGCAAGGGGGAGTGGCCCTGSCCAACGTGACTTGTCTCC 1500
Qy      ||||| 1501 AACCTGAGCCCGCAGCAACAGAGGCTCTCCCTTGATCAAGCCGCTGTGAATGACAG 1560
Db      ||||| 1501 AACCTGAGCCCGCAGCAACAGAGGCTCTCCCTTGATCAAGCCGCTGTGAATGACAG 1560
Qy      ||||| 1561 GCCTGCCCCCGAGCGAGATGAGAGTGAAGGCCCGAGAGAAAGATTTCCCACTGATAAA 1620
Db      ||||| 1561 GCCTGCCCCCGAGCGACGATGAGATGAGAGGCCCGAGAGAAAGATTTCCCACTGATAAA 1620
Qy      ||||| 1621 GGGTACTTATATGCTTAAGGAAGTGTCTACACCGAGCGAACTATCTGAAGATCTCGAA 1680
Db      ||||| 1621 GGGTACTTATATGCTTAAGGAAGTGTCTACACCGAGCGAACTATCTGAAGATCTINGAA 1680
Qy      ||||| 1681 GTTATCACTTGTGTTTCAAGACACAGTGAACAAAGAGAGCCCATGCGGAGAGACTG 1740
Db      ||||| 1681 GTTATCACTTGTGTTTCAAGACACAGTGAACAAAGAGAGCCCATGCGGAGAGACTG 1740
Qy      ||||| 1741 AAAAGCTATATTTCCGAAATTTTGAACCTTTGACAAATTTCTAATCTAATTTTCTCAAG 1800
Db      ||||| 1741 AAAAGCTATATTTCCGAAATTTTGAACCTTTGACAAATTTCTAATCTAATTTTCTCAAG 1800
Qy      ||||| 1801 GAAATTTGACCAAGACTTCCCTGTGAGAGAGCCGCTCAATATGCCAATTCAGAGATTAC 1860
Db      ||||| 1801 GAAATTTGACCAAGACTTCCCTGTGAGAGAGCCGCTCAATATGCCAATTCAGAGATTAC 1860
Qy      ||||| 1861 CAAAGATGCGGAGATGCTATGCTGAAGAACATTCAGGGCATGAAGCACTGGCGGCTCAC 1920
Db      ||||| 1861 CAAAGATGCGGAGATGCTATGCTGAAGAACATTCAGGGCATGAAGCACTGGCGGCTCAC 1920
Qy      ||||| 1921 CTGTGAGACACAGCGAGGCTTTGAGAGGCTCTGAGAGATGGAATCAAGAGCTTCCGCGG 1980
Db      ||||| 1921 CTGTGAGACACAGCGAGGCTTTGAGAGGCTCTGAGAGATGGAATCAAGAGCTTCCGCGG 1980
Qy      ||||| 1981 CTGAGAGACTTCTGAGAGACTTTGAGCTGAGAGAGGTGTGTTACTTACCGCTCAACACC 2040
Db      ||||| 1981 CTGAGAGACTTCTGAGAGACTTTGAGCTGAGAGAGGTGTGTTACTTACCGCTCAACACC 2040
Qy      ||||| 2041 TTCTCTGCGGCGCACTGCAACGCGCTCATGCACTAACAGAGTCTGAGCGGCTGTGC 2100
Db      ||||| 2041 TTCTCTGCGGCGCACTGCAACGCGCTCATGCACTAACAGAGTCTGAGCGGCTGTGC 2100
Qy      ||||| 2101 AAAACACACCCGCGGAGCCACGCGCACTTCAAGGACTGCGGACCGCTTTGCAAGATC 2160
Db      ||||| 2101 AAAACACACCCGCGGAGCCACGCGCACTTCAAGGACTGCGGACCGCTTTGCAAGATC 2160
Qy      ||||| 2161 ACGAGATGATGAGCAGCTCCACGAGTCAAGATGATCAAGATGGAAGATTTTCCAGAGCTG 2220
Db      ||||| 2161 ACGAGATGATGAGCAGCTCCACGAGTCAAGATGATCAAGATGGAAGATTTTCCAGAGCTG 2220
Qy      ||||| 2221 CACGAATCTCAAGAAAGATTTGATTTGAGATTGACAAATCTTGTGTTCCGGAAAGGAGTTC 2280
Db      ||||| 2221 CACGAATCTCAAGAAAGATTTGATTTGAGATTGACAAATCTTGTGTTCCGGAAAGGAGTTC 2280
Qy      ||||| 2281 ATCCGTGTGGGAGGCTTCAAGAGCTCTCGGGAGAGGGCTCCAGAGCGCATGTTCTTC 2340
Db      ||||| 2281 ATCCGTGTGGGAGGCTTCAAGAGCTCTCGGGAGAGGGCTCCAGAGCGCATGTTCTTC 2340
Qy      ||||| 2341 CTGTTCAACGAGCGTCTGCTATTAACAAGCCGGGGCTGACGAGCTCAATCGATGTTTAA 2400
Db      ||||| 2341 CTGTTCAACGAGCGTCTGCTATTAACAAGCCGGGGCTGACGAGCTCAATCGATGTTTAA 2400
Qy      ||||| 2401 GTTCACGGGACGCTCCGCTCTATGAGCATGAGATGAGAGACGAGAGCGAGTGGGG 2460
Db      ||||| 2401 GTTCACGGGACGCTCCGCTCTATGAGCATGAGATGAGAGACGAGAGCGAGTGGGG 2460
Qy      ||||| 2461 GTGCCCACTGCTGACCTTCCGGGGCCAGCGGAGTCAATCATGTGAGCCGCAATTTCT 2520
Db      ||||| 2461 GTGCCCACTGCTGACCTTCCGGGGCCAGCGGAGTCAATCATGTGAGCCGCAATTTCT 2520
Qy      ||||| 2521 CGGTCCGAGATGAGAGATGAGGTTTGAAGCATTCAGATGGCATTTGACCTGCGGAGAA 2580
Db      ||||| 2521 CGGTCCGAGATGAGAGATGAGGTTTGAAGCATTCAGATGGCATTTGACCTGCGGAGAA 2580
Qy      ||||| 2581 AGCAGAGCCCGCCCTTGAAGTTCTGTGCGCAGACGCCCCCTGACAAAGTCCCTGAT 2640
Db      ||||| 2581 AGCAGAGCCCGCCCTTGAAGTTCTGTGCGCAGACGCCCCCTGACAAAGTCCCTGAT 2640
Qy      ||||| 2641 GAAAGCCACCGCGGCTGACAGAGATGAGAGTGAAGGCTTCCGCGCAAGTGTGCTG 2700
Db      ||||| 2641 GAAAGCCACCGCGGCTGACAGAGATGAGAGTGAAGGCTTCCGCGCAAGTGTGCTG 2700
```

QY	2701	GAGGCGCAGGCCCCCGCACCGGGCAACAATGTGCACTGTGCGCACCGCAACC	2766
Db	2701	GAGCGCCAGGCCCCCGCACCGGGCAACAATGTGCACTGTGCGCACCGCAACC	2766
QY	2761	AGCGTCTCCATGTGTGAATTCCAGCATCGCAGTGGAGAATCAGTTGTCGGAAACCTGCTG	2820
Db	2761	AGCGTCTCCATGTGTGAATTCCAGCATCGCAGTGGAGAATCAGTTGTCGGAAACCTGCTG	2820
QY	2821	AGGAAATTCAAAAACAGCAAGGGGTGGCAAGCTGTGGGTGTTCACAACTTCTGC	2880
Db	2821	AGGAAATTCAAAAACAGCAAGGGGTGGCAAGCTGTGGGTGTTCACAACTTCTGC	2880
QY	2881	CTGTCTCTCTCAAAATCACACGAGGACATATCCCTTTGCGAGCTGCTCTGTGCGC	2940
Db	2881	CTGTCTCTCTCAAAATCACACGAGGACATATCCCTTTGCGAGCTGCTCTGTGCGC	2940
QY	2941	TACTGCGTCACCATCCCTCTGTAGTCCGAGAACATCCAGAAAGACTACGTGTCAAGCTG	3000
Db	2941	TACTGCGTCACCATCCCTCTGTAGTCCGAGAACATCCAGAAAGACTACGTGTCAAGCTG	3000
QY	3001	CACCTCAAGTCCCAAGTCTACTACTTCAAGGGCGGAABCGAGTACAGTTGGAAAGGTGG	3060
Db	3001	CACCTCAAGTCCCAAGTCTACTACTTCAAGGGCGGAABCGAGTACAGTTGGAAAGGTGG	3060
QY	3061	ATGGAAGTATCCGCGAGTCCACCAAGCTCTGCTCGACACCCCAAGTGTGAGCCACAAA	3120
Db	3061	ATGGAAGTATCCGCGAGTCCACCAAGCTCTGCTCGACACCCCAAGTGTGAGCCACAAA	3120
QY	3121	GAGTCTCTGTGTGTAT	3135
Db	3121	GAGTCTCTGTGTGTAT	3135

RESULT 4	3147 bp	DNA	linear	GSS 15-DEC-2001
LOCUS AY406169				
DEFINITION Mus musculus FARP1 gene, VIRTUAL TRANSCRIPT, partial sequence,				
ACCESSION AY406169				
VERSION AY406169.1				
KEYWORDS GSS.				
SOURCE Mus musculus (house mouse)				
ORGANISM Mus musculus				
REFERENCE Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidae; Muridae; Murinae; Mus.				
REFERENCE 1 (bases 1 to 3147)				
AUTHORS Clark,A.G., Glatowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferritera,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.T., Adams,M.D. and Cargill,M.				
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios				
JOURNAL Science 302 (5652), 1960-1963 (2003)				
PUBMED 14671302				
REFERENCE 2 (bases 1 to 3147)				
AUTHORS Clark,A.G., Glatowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferritera,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.T., Adams,M.D. and Cargill,M.				
TITLE Direct Submission				
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA				
COMMENT This sequence was made by sequencing genomic exons and ordering them based on alignment.				
FEATURES				
SOURCE				
1..3147				
/organism="Mus musculus"				
/mol_type="genomic DNA"				
/db_xref="taxon:10090"				
<1..3147				
/gene="FARP1"				
gene				

[illegible]

Dh 961 GAACCTTAAGCCAAAGCCAAAGCCTGTTCTTCAGCCGAGGTCGTCTTCGCGGTTAAT 1020
Qy 1021 GGTTCGACTCAGAAAGCAGGTTCTCGACTATGTTAAAGAAAGAGACATAAGAAAGTTCAG 1080
Db 1021 GGTTCGACTCAGAAAGCAGGTTCTTCGATTAATGTTAAAGAAAGCGCCACAAGAAAGTTCAG 1080
Qy 1081 TTTGAAGAGAGACAGCAAGATTTCAATTCTATCCGAGCCTTGCTTCAAGCCTTACAGAA 1140
Db 1081 TTTGAAGAGAGACAGCAAGATTCACATCCAGAGAGCCTGATCTCAAGCCCAACAGCA 1140
Qy 1141 CTGAATTCGGAAGTGTCTGAGCAGTCTCAGCAGAGACACAGCCTTACATTGGAGAAGT 1200
Db 1141 CCGAATTCAGAAATGTCCAAAACAGTCTCCACAGAGCGCAAGCCTTACGTTGGAGAAGT 1200
Qy 1201 GCCGAATTCACAGGGGGCCAGAGCTCCGCGGAGAGAAAGAACCGAAGTTCCCGCGGG 1260
Db 1201 ACCGAGTCTCCGCGGTGAGCAGAGCTCCAGCAAGCGAAGAAACAAAGCTGTACCTTG 1260
Qy 1261 GAGCCGAGGTCGACCCGAGCCCTGCGCGGAGAGAAAGCCCGGAGTAAACAAGCAGCG 1320
Db 1261 GAGTTGGAGCTTACCAAGCCCCCGCTTCCTTAAGAGCCTTCAGAGAGCAAAAGCTGCA 1320
Qy 1321 GACGAGCCGCTTCGCGCGCCACAGAGAA-----GAGAGAGAGTCTGTTAAGATAG 1374
Db 1321 GATGGAACCAACGTTGGTCGCTGAGAGAAAGAGAGAGAGAGAGAGTGGCAAGACGCG 1380
Qy 1375 ACCCAGCAGATTAACCTCAGCCCCCGCAGACCAAGCAGAGCTCCTGACTGGCAATCT 1434
Db 1381 ATCCGCGCAAGTAACCTCAGCCCCCGCAGACCAAGCAGAGCTCCTGATGGTACCTCT 1440
Qy 1435 CACCTTTCGAGCTGTCTGGAATCGCAGAGGGGAGTGGCCCTGCAACGTCGACTTG 1494
Db 1441 CACCTTTCAGAGCTGTCTGATCACTCAGAGGAGAGAGCGCCCAAGCAGTGAACCTG 1500
Qy 1495 TCTCCCAACCTGAGCCCGACACCAAGCAGAGCCTTCTCTTATCAGCCGCTGTGAAT 1554
Db 1501 TCTCCCAACCTGAGTCTGACCAACAGCAGGCTCTCTCTGATCAGCCGCTCTGAAC 1560
Qy 1555 GACCAAGCCTCCCGCGAGAGAGAGATGAGAGATGAGGGCCGAGAGAAAGATCCCACT 1614
Db 1561 GACCAAGCCTGCCCAAGAGACCGACAGAGAAAGAGGCGGAGAAAGATTCACCAAC 1620
Qy 1615 GATPAAGCTACTTCTTATGAGAAAGTGTACCAACGAGGAAACATATCTGAAGAT 1674
Db 1621 GACAAAGCGTACTCATTTGCTTAAGAAAGTCTCAACACAGAGGCACTGATCTGAAGAC 1680
Qy 1675 CTGAAGTATCACTTCGTGTTTCAGAGCAGTGAAGCAAGAGACCGCATGCGGAA 1734
Db 1681 CTGAGGTCTATGCTTCGTGTTTCAGAGCAGTGAAGAGACCTCATGCGGAA 1740
Qy 1735 GCACTGAAAAGTCTCATATTCGGAATTTGAACTTTGCAAAATTTCTACTAATTT 1794
Db 1741 GCCTTGAAGGTCTCATATTCGAAATTTGAACTTTGCAAAATTTCTACTAATTT 1800
Qy 1795 CTCAAGAAATTGAGCAAGACTTGCCTGTGGAAAGCGCTCAATGCGCAATCA-- 1852
Db 1801 CTGAAGAAATTGAGCAAGACTTGCCTGTGGAAAGCGCTCAATGCGCAATCAAG 1860
Qy 1853 -GAGATTACCAAGAAATCGCGAGTGTCTATGTAAGAACATTGAGGCGATGAAGACTG 1911
Db 1861 GAGATTACCAAGAAATCGGAGAGTGTCTATGTAAGAAACATTGAGGCGATGAAGACTG 1920
Qy 1912 GCGGCTCACTGTGGAAGCAGCAGAGGCTTGGAGGCTTGGAGAAATGAATCAAGAGC 1971
Db 1921 GCGGCGCACTGTGGAAGCAGTGAAGGCTTGGAGGCTTGGAGAACTTCAATCAAGAGC 1980
Qy 1972 TCCCGGCGGTGAGAACTTCTGAGAGACTTGAAGTGAAGAGGTGTTACTTACCG 2031
Db 1981 TCGCGGCGGTGAGAACTTCTGCGAGAGCTTGAAGTGAAGAGGTGTTACTTACCG 2040
Qy 2032 CTCAACACCTTCTCTGCGGCGCACTGACCGGCTCATGCACTAAGCAGGCTCTGAG 2091
Db 2041 CTCAACACCTTCTCTGCGGCGGCTGACCGGCTCATGCACTAAGCAGTGTCTGAG 2100

Qy 2092 CGGCTGTGCAAAACCAACCCGCGAGCCAGCGGCACTTCAGGGAATGCGGACCGCTTG 2151
Db 2101 AGGCTGTGCAAGACCAACCAACCGCGGCACTTCAGGGAATGCGGAGCTGCGAGCTG 2160
Qy 2152 GCAGAGATCACGAGATGCTGAGCAGCTCAAGGTCAGATGATCAAGATGAAGAAATTC 2211
Db 2161 GCGAGATCACAGAGATGCTGAGCAGCTGATGATGATCAAGATGAAGAAATTC 2220
Qy 2212 CAGAACTGCAAGAACTCAAGAAATTTGATTTGCAATTTGATTTGATTTGATTTG 2271
Db 2221 CAGAACTGCAAGAACTCAAGAAATTTGATTTGATTTGATTTGATTTGATTTG 2280
Qy 2272 AGGAGTTCATCTGCTGAGGAGCTCAGCAAGCTTCGAGGAGAGGAGCTCAGCAGCGC 2331
Db 2281 AGGAGTTCATCTGCTGAGGAGCTCAGCAAGCTTCGAGGAGAGGAGCTCAGCAGCGC 2340
Qy 2332 ATGTTCTTCTGTTCAACGAGCTCTGTATACAGAGCGGAGGCTGACGAGCTTCAAT 2391
Db 2341 ATGTTCTTCTGTTCAACGAGTCTGTATACAGAGCGGAGGCTGACGAGCTTCAAT 2400
Qy 2392 CAGTTTAAATCCAGGAGAGCTCCGCTCTATGAGATGATGATGAGAGAGAGAGC 2451
Db 2401 CAGTTTAAATCCAGGAGAGAGCTCCCACTATGAGATGATGATGAGAGAGAGAG 2460
Qy 2452 GAGTGGGAGGAGCCCACTGCTGACCTCCGAGGCGCAGGCAAGTCAATCGTGGCC 2511
Db 2461 GAGTGGGAGGAGCCCACTGCTGACCTCCGAGGCGCAGGCAAGTCAATCGTGGCT 2520
Qy 2512 GCCAGTTCGTGTCGAGATGAGAGAGTGGTTGAGAGATCATCAATGAGCAATTCAG 2571
Db 2521 GCCAGTTCGTGTCGAGATGAGAGAGTGGTTGAGAGATCATCAATGAGCAATTCAG 2580
Qy 2572 GCGAGAAAG 2631
Db 2581 GCGAGAAAG 2640
Qy 2632 TCCCTGATGAAG 2691
Db 2641 TCCCTGATGAAG 2700
Qy 2692 ACATGCTGAG 2751
Db 2701 ACATGCTGAG 2760
Qy 2752 CGCAACACAG 2811
Db 2761 CGCAACACAG 2820
Qy 2812 AACCTGCTGAG 2871
Db 2821 AACCTGCTGAG 2880
Qy 2872 AACCTGCTGAG 2931
Db 2881 AACCTGCTGAG 2940
Qy 2932 CTGCTGAGTACTGCTGACATCCCTCTGAGTCCGAGAAACATCCAGAAAGACTACG 2991
Db 2941 CTGCTGAGTACTGCTGACATCCCTCTGAGTCCGAGAAACATCCAGAAAGACTACG 3000
Qy 2992 TTCAAGCTGACTTCAAGTCCCAAGCTCTACTTCAAGGAGAGAGAGAGAGAGTTC 3051
Db 3001 TTCAAGCTGACTTCAAGTCCCAAGCTCTACTTCAAGGAGAGAGAGAGAGAGTTC 3060
Qy 3052 GAAAGTGAATGAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3111
Db 3061 GAAAGTGAATGAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3120
Qy 3112 AGCACAAGAGTCTTGTGTA 3134
Db 3121 AGCACAAGAGTCTTGTGTA 3143

[illegible]

Guerninich, S., Hirokawa, N., Jackson, I. J., Jarvis, E. D., Kanaï, A., Kawaji, H., Kanaawa, Y., Kedzierski, R. M., King, B. L., Konagaya, A., Kurochkin, I. V., Lee, Y., Lenhard, B., Lyons, P. A., Maglott, D. R., Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numata, K., Okado, T., Pavan, M. J., Pereira, G., Pesole, G., Petrovsky, N., Pflaj, R., Pontius, J. U., Qi, D., Ramachandran, S., Ravasi, T., Reed, J. C., Reed, D. J., Reid, J., Ring, B. Z., Ringdahl, M., Sandelin, A., Schneider, C., Sempke, C. A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M. S., Teasdale, R. D., Tomita, M., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watnabe, Y., Wells, C., Wilming, L. G., Wysshaw-Borja, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hayazane-Kishikawa, T., Kono, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, I., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imocani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shimagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E. S., Rogers, J., Birney, E. and Hayashizaki, Y.

PANTOM Consortium

CONSRM
TITLE
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL
Nature 420 (6915), 563-573 (2002)
PUBMED
12466851
REFERENCE
6

Carninci, P., Kasukawa, T., Katayama, S., Gough, J., Frith, M. C., Maeda, N., Oyama, R., Ravasi, T., Lenhard, B., Wells, C., Kodzius, R., Shimokawa, K., Bajic, V. B., Bremner, S. E., Batalov, S., Forrest, A. R., Zavolan, M., Davis, M. J., Wilming, L. G., Aldin, V., Allen, J. E., Ambesi-Imbimbato, A., Apweiler, R., Auzaiya, R. N., Bailey, T. L., Bansal, M., Baxter, L., Belsel, K. W., Bersano, T., Bonho, H., Chalk, A. M., Chiu, K. P., Choudhary, V., Christoffels, A., Clutterbuck, D. R., Crowe, M. L., Dalla, E., Dalrymple, B. P., de Bonis, B., Della Gatta, G., di Bernardo, D., Down, T., Engstrom, P., Engstlin, M., Faulkner, G., Fletcher, C. F., Fukushima, T., Furuno, M., Futaki, S., Gaitoldi, M., Georgii-Hemming, P., Gingeras, T. R., Gojobori, T., Green, R. E., Gustcinich, S., Harbers, M., Hayashi, Y., Hensch, T. K., Hirokawa, N., Hill, D., Hummel, L., Iacono, M., Ikeo, K., Iwama, A., Ishikawa, T., Jakt, M., Kanpin, A., Katoh, M., Kawasawa, Y., Kelso, J., Kitamura, H., Kitano, H., Kollias, G., Kristiansen, S. P., Kruger, A., Kummerfeld, S. K., Kurochkin, I. V., Lareau, L. F., Lazarevic, D., Lipovich, L., Liu, J., Liuni, S., McWilliam, S., Madan, Babu, M., Madera, M., Marchionni, L., Matsuda, H., Matsuzawa, S., Miki, H., Mignone, P., Miyake, S., Morris, K., Mottagui-Tabar, S., Mulder, N., Nakano, N., Nakachi, H., Ng, P., Nilsson, R., Nishiguchi, S., Nishikawa, S., Nori, F., Ohara, O., Okazaki, Y., Orlando, V., Pang, R. C., Pavan, W. J., Pavese, G., Pesole, G., Petrovsky, N., Piazza, S., Reed, J., Reid, J. F., Ring, B. Z., Ringdahl, M., Roost, B., Ruan, Y., Salzberg, S. L., Sandelin, A., Schneider, C., Schonbach, C., Sekiguchi, K., Sempke, C. A., Seno, S., Sessa, L., Sheng, Y., Shibata, Y., Shimada, H., Shimada, K., Silva, D., Sinclair, B., Sperling, S., Stupka, E., Sugitara, K., Sultana, R., Takenaka, Y., Taki, K., Tamoya, K., Tan, S. L., Tang, S., Taylor, M. S., Tegner, J., Teichmann, S. A., Ueda, H., van Nimwegen, E., Verardo, R., Wei, C. L., Ye, J. K., Yamashita, H., Zaborov, E., Zhu, S., Zimmer, A., Hilde, W., Bult, C., Grimmond, S. M., Teasdale, R. D., Liu, E. T., Brusic, V., Quackenbush, J., Wahlestedt, C., Matlick, J. S., Hume, D. A., Kai, C., Sasaki, D., Tomaru, Y., Fukuda, S., Kanamori, K., Katayama, T., Suzuki, M., Aoki, J., Arakawa, T., Iida, J., Imamura, K., Itoh, M., Kato, T., Kawaji, H., Kawagashira, N., Kawashima, T., Kojima, M., Kondo, S., Kono, H., Nakano, K., Niimura, N., Nishio, T., Okada, M., Plesky, C., Shibata, K., Shiraki, T., Suzuki, S., Tagami, M., Waki, K., Wachihi, A., Okamura-Ono, Y., Suzuki, H., Kawai, J. and Hayashizaki, Y.

PANTOM Consortium

CONSRM
TITLE
The transcriptome landscape of the mammalian genome
JOURNAL
Science 309 (5740), 1559-1563 (2005)
PUBMED
16141072
REFERENCE
7

Katayama, S., Tomaru, Y., Kasukawa, T., Waki, K., Nakashima, M., Nakamura, M., Nishida, H., Yap, C. C., Suzuki, M., Kawai, J., Suzuki, H., Carninci, P., Hayashizaki, Y., Wells, C., Frith, M., Ravasi, T., Pang, K. C., Hallinan, J., Matlick, J., Hume, D. A., Lipovich, L., Batalov, S., Engstrom, P. G., Mizuno, Y., Feghli, M. A., Sandelin, A., Chalk, A. M., Mottagui-Tabar, S., Liang, Z., Lenhard, B. and

QY	1279	GGCCCTGGCGCGAGGAGAAAGCCCGCGGGGTAAACAAGAGGGGGAGACGAGCGCCGCTCGGCG	1338
Db	1419	CTCCCTTTGGATCCCGACAGGCTCCCGTATCAAGAGCAACAGAGAGAGAGAGACGCG	1478
QY	1339	CCCAACGAGGAAGAGAGAGAGGTCCGTTAAGTATGAGACCCAGACAGTAAACTCAGCCC	1398
Db	1479	ACGA--TCATCATCTGACGCGGCCAGACACAGTGGGCCAATCTCCCTGGACCCCGGTG	1536
QY	1399	CCGACGCCAAGCAACAGGCTCCCTGACTGACAGTCTCACTCTTCCGAGCTGTCTGTAAAC	1458
Db	1537	CTCCGGCCCTGGTCCAGGCTTTTCTATGATGATGTCTCAGCCTTCTCCCTCCAGCCTGAAG	1596
QY	1459	TCCGAGGGGGGAGTGGCCCCCTGGCCAACGTGACTTGTCTCCCAACCTGAGCCCCGACAC	1518
Db	1597	AGCCACTGACCTGTGCTGTGCTTGAGCTTCAGGCGCG-----CTGAGTACAGCTGAG	1647
QY	1519	AAGCAGGCGCTCTCCCTTGATCAGCCCGCTGCTGAATGACACAGGCTCGCCCCGAGCGAC	1578
Db	1648	CAGGGTGCATCCCTGTGTCTAGCCCTGTGTCTCACTGGCCGTGGTACAGCCAGATGAC	1707
QY	1579	GATGAGATGAGGGCCGAGGAGAGATTTCCCACTGATAAAGCTACTTCATAGCTAAG	1638
Db	1708	AACCAAGAAAGACAGAAACACACAGACATGCCAGAAAGATGAGGCTTATTCATAGCCAA	1767
QY	1639	GAAGTGTCTACCAACCGAGCAACATATCTGAAAGATCTCGAAAGTTATCACTTCTGGTTT	1698
Db	1768	GAGATTTCTCGCTACGAAACGAACCTATCTGAAGGATTTGAAAGTTATCACTGTGTGGTTC	1827
QY	1699	CAGAGCAGTGAAGCAAGAGAGACGCACATGCGGAAAGCATGAAAGTCAATATTCGCG	1758
Db	1828	CGAGGCGTGTGATTCAGAGAGAGGACCATGCTGACAGCCTGATGAGCCCTGTTTCTCC	1887
QY	1759	AATTTTGAACCTTTGCAACAAATTTCTATCTAATTTCTCAAGAAATTGACCAACGACTT	1818
Db	1888	AAACATTTGATCAGATCTACAGATTCACAGAGGCTTCTTCAAGATGGAACAGAGGTG	1947
QY	1819	GCCCTGTGGGAAGGCGCGCTCAAAATGCCA---AATCAGATTTACCAAGATGGCGAT	1875
Db	1948	GCACCTCTGGGAGGGGCCCTCCAGTGCCCACTTAAAGGTGATCTACACAGCGAATGGGGAC	2007
QY	1876	GTCATGTGGAAGAACATTCAAGGGCATGAAGACCTCGGGCGCTCACCTGTGGAAGACAC	1935
Db	2008	ATCTCTCTCAGAAACATGCTGTCAGTTAAAGAAATTTACTAGCTACTTCAAAAGACGAT	2067
QY	1936	GAGGCTTTGAGGCGCTCGAGATGGAATCAAGAGCTCCCGCGCTGGAACCTTCTGC	1995
Db	2068	GAGGTCTCAACAGAACTGGAAAGGCCACAAACACTGTAAAAAGCTGAGAGGACTCTAC	2127
QY	1996	AGAGACTTTGAGCTGACAGAGGTGTGTTACTTACCGCTCAACACCTTCTCTCGGCGCA	2055
Db	2128	AAAGAGTTTGAAGCTCAAAAAGTGTGCTACCTGCTCAACACATTTCTCTGAAGCC	2187
QY	2056	CTGCAACCGGCTCATGCACTACAGAGGTCTCTGAGCGGCTGTGCAACACACCCGCG	2115
Db	2188	GTCAGAGGCTAGTCACTACCTGCTGTGCTGAGCGCGCTGTGCTCACTACTCTCT	2247
QY	2116	AGCCACGCGCACTTCAGGGACTGCGACGCGCTTTTGGCAGAGATCAACGAGATGTGGCA	2157
Db	2248	GGGCAACGGCACTATGCGGCACTGCCATGAGGCACTGAAGGCCATCAAGAAAGTACACAC	2307
QY	2176	CAGCTCCACGGTACGATGATCAAGATGAGAAATTTCCAGAACTGCAAGAACTCAAGAA	2235
Db	2308	GAGCTCCAGCAAAAGCTTACCCGCGTGGAAAACTCAGAAATTTGACGGAAGCTACAGGA	2367
QY	2236	GATTTGATTTGGCATTTGACAAATCTTGTGTTTCCGGGAAGGAGTTCAATCCGCTGTGGCAGC	2295
Db	2368	GACCTGTGTGGTGTAGAAACCTCATTTGCTCTGTGAGAGGAGTTTATCCGTAGAGGCGTGC	2427
QY	2296	CTCAGCAAGCTCTCGGGGAAAGGGGCTCAGAGCGCATGTTCTTCTGTTCAAGCAGCTC	2355
Db	2428	CTGCACAAAGCTCAACCAAGAGGCGCTGCAACAGAGAGATGTTTTTCTGTCTCAGATATG	2487
QY	2356	CTGCTATACACGAGCGGGGCTGACGGCTTCATTCAGTTTAAATTCACAGGCGAGCTC	2415

Db	2448	TTACTGTTATACAGCAAAAGTGTACACAGAGCCAGTCAATTTCCGGATCCGGTCTCTT	2547
Qy	2416	CCGCTCTATAGGCATGACGATTTGAGAGAGCGAAGCAGATGGGGGGTCCCCATGCTCG	2475
Db	2548	CCACTCCGTTGGCATGTGTTGAGAGAAAGTGAAGATGAATGATCTGTTCTTCAATTCCTTC	2607
Qy	2476	ACCCTTCGGGGGCCACGGCCAGTCCATCATCTGTGGCCGCCAGTTTTCGGTCCGAAATGAG	2535
Db	2608	ACCAATCTATATGAGCTCAGAAAACAATTTGTGTATCACCCAGCACTCGGCTTAGAAAAAGAA	2667
Qy	2536	AAGTGGGTTTGAAGACATTCAGATGGCCATTGACCTGGCGGAGAAAGACAGACGCCGCC	2595
Db	2668	AAGTGGATGACAGGACCTGAAATGACGCAATCCAGACAGCCAAGCTATCGGTACTCACCC	2727
Qy	2596	CCTGAGTTCCTGGCCAGACGCCCTTCGACAAACAAGTCCCTCTGATGAAGCACCGCGAGT	2655
Db	2728	CCAGTGGCTGTGG---GAGGCCGGGTGTATATCTGTATCCCTATAGCTTTCATGAAGTGC	2784
Qy	2656	GACCAAGATGACAGAGATGACCTTAGAGCGCTCGCGCACATGCTGGAGCCGACGCCCG	2715
Db	2785	TCTCTGGAAAGATCAGAAAGATGTCTGAGAAACCGGGGGTCTCCCTGAGGGGAAACAGCAG	2844
Qy	2716	CACCGCGGCAACACAAATGTGACGTGTGCGTGACCGGCAACACACAGCTTCATGATGTG	2775
Db	2845	CACGGGGCAATACAAACATGATGTGTGCTGGTACCGTAATACAGTGTGTCCAGAGCA	2904
Qy	2776	GACTTTCAGCATCGCAGTGGAGAAATCAGTGTCTTGAAACCTGCTGAGAAATTCMAAAC	2835
Db	2905	GACCAACAGTGAAGCTGTGTAGAACACAGCTTTCAGGATATCTGCTGAGAAAGTTCAAGAC	2964
Qy	2836	AGCAACGGGTGGCAGAAAGCTGTGGTGTGTTCACAAACTTTCGCTGTTCTTACAA	2895
Db	2965	AGTATGTGCTGGCAGAAAGCTGTGGTGTGTTCACAACTTTCGCTGTTCTTACAA	3024
Qy	2896	TCAACACAGGACATCAATCCCTTGGCCAGCCCTGCTGCTCGGCTACTCGCTACATC	2955
Db	3025	ACACATCAGGATGACTACCCCTTGGCCAGCTCCACATATTTGGTATACAGGTGAGCTC	3084
Qy	2956	CCCTCTGAGTCCGAGAAATCACAAGAACTACGTGTTCAAGCTGCACCTTCAAGTCCAC	3015
Db	3085	CCCAAGGAGGCTGACAGCATCAACAAAGATATGTTCTTCAAGCTCCAAATTCATCTCAC	3144
Qy	3016	GTCTACTACTTTCAGGGCGGAAAGCGAGTACAGTTTGAAAGGTGATGAAATGATCCGC	3075
Db	3145	GTCTACTTCTTCCGGGCTGAGAGCAATGACATTTGAAAGGTGATGAGCATCAAA	3204
Qy	3076	AGTGCACACAGCTCTGCTGCGCACCCCAAGTGTGAGCCACAAAGAGTCTTT	3129
Db	3205	AGGGCCAGCACTCACAGGAGAGACCCCAAGTTTCACTACAGAGCTGCTCAAT	3258

RESULT 6

AK050860

LOCUS

DEFINITION

AK050860.1 GI:26094164

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,

3874 bp

mRNA

linear

HTC 02-SEP-2005

Mus musculus 9 days embryo whole body cDNA, RIKEN full-length

enriched library, clone:D030026M03 Product:weakly similar to CDEP

[Homo sapiens], full insert sequence.

AK050860

AK050860.1 GI:26094164

HTC; CAP trapper.

Mus musculus (house mouse)

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

Sciurognathi; Muridae; Muridae; Murinae; Mus.

1

Carninci, P. and Hayashizaki, Y.

High-efficiency full-length cDNA cloning

Meth. Enzymol. 303, 19-44 (1999)

10349636

2

TITLE	Itch,M., Komo,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)
PUBMED	11042159
REFERENCE	
AUTHORS	3 Shibata,K., Itch,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Kono,H., Akiyama,J., Nishi,K., Kitasuna,T., Tashiro,H., Itch,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishino,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kaishiwagi,K., Fujiwaka,E., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watabiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y. RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)
PUBMED	11076861
REFERENCE	
AUTHORS	4
TITLE	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
JOURNAL	Functional annotation of a full-length mouse cDNA collection
REFERENCE	Nature 409, 685-690 (2001)
AUTHORS	5
TITLE	The PANTOM Consortium, the RIKEN Genome Exploration Research Group Phase I and II Team.
JOURNAL	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
REFERENCE	Nature 420, 563-573 (2002)
AUTHORS	6
TITLE	RIKEN Genome Exploration Research Group, Genome Science Group (Genome Network Core Team) and the FANTOM Consortium.
JOURNAL	Antisense transcription in the Mammalian Transcriptome
REFERENCE	Science 309, 1564-1566 (2005)
AUTHORS	7
TITLE	The PANTOM Consortium, Riken Genome Exploration Research Group and Genome Science Group (Genome Network Project Core Group).
JOURNAL	The Transcriptional Landscape of the Mammalian Genome
REFERENCE	Science 309, 1559-1563 (2005)
AUTHORS	8 (bases 1 to 3874)
TITLE	Adachi,J., Aizawa,K., Akimura,T., Arikawa,T., Bono,H., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,M., Hayashida,K., Hayatsu,N., Hiramoto,K., Hirooka,T., Hirozane,T., Horii,P., Imocani,K., Ishii,Y., Itch,M., Kagawa,I., Kanakawa,T., Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Kono,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,Y., Nishi,K., Nomura,K., Numasaki,R., Ohno,M., Ohsato,N., Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Shibata,K., Shingawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takai-Akahira,S., Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.
JOURNAL	Direct Submission
REFERENCE	
AUTHORS	Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsukumi-ku, Yokohama, Kanagawa, 230-0045, Japan (E-mail:genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)
COMMENT	cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details. URL:http://genome.gsc.riken.jp/ URL:http://fantom.gsc.riken.jp/. Location/Qualifiers
FEATURES	1..3874
SOURCE	organism="Mus musculus" mol_type="RNA" strain="C57BL/6J" db_xref="FANTOM_DB:D030026M03"

Query Match	Similarity	Score	DB	Length
Best Local Match	59.1%;	1007.6;	DB 6;	3874;
Matches 1843;	Conservative	0;	Mismatches 1244;	Indels 29; Gaps 6;

misc_feature	note="putative	evidence: FASTA, 53.7kID, 99.7klength, match=3125"
1. .3874		
weakly similar to CDBP [Homo sapiens] (SPTRIQ9Y4F1,		

ORIGIN	/db_xref="taxon:10090"	/clone="D030026M03"	/issue_type="whole body"	/clone_id="RIKEN full-length enriched mouse cDNA library"
Query Match	32.1%;	Score 1007.6;	DB 6;	Length 3874;
Best Local Match	59.1%;	Pred. No. 5.7e-255;		
Matches 1843;	Conservative	0;	Mismatches 1244;	Indels 29; Gaps 6;

Query	Subject	Score	DB	Length
22	CCGACCCAGGATACGACCTGGGGGGCCCGGAAATTCGGGGATCAGTACCTTGGAACGT	81		
Db	103	CCAACTTCAGGGACACGCTGGGGGCGACGCCCATTTGAGTGAACCTTGAGGCA	162	
Qy	82	GGACAGAGCGCGCCCCCAACACTTTCAGGAAATCTGCTGCATCAAAATCAGATCGTG	141	
Db	163	GAGCAGAGCCTGTGACCCAGGATCAGAGAGAGACATGGCATCAGATGAAGTGGTG	222	
Qy	142	GATGACACCCAGGAGGCAATTGAAATTCACAAAGAGCTCCTGGAGAGTCTGTGAT	201	
Db	223	GACGACCTGATGATTTGATGATGAGCCGAGAGTGAACGCGGACAGTCTTACTGACT	282	
Qy	202	GCAGTTTGCAACCACTTCACCTCGTGGAGAGTACTATTATTGGCTCTGATCTCTAT	261	
Db	283	CAATGTGAGAGCATTTGAAATCTCATTTGAATGTATCACTTTGGTCTGAGATTCAGAT	342	
Qy	262	CACAAAAAGATCACGCTGTGGCTGATCTCCTAAACCCATTGTGAAACAGATTGAAG	321	
Db	343	GTCCAGTCTACTGATTTGGCTTGAACCTATGAAACCATCACTTAGGCAAGTACGAAG	402	
Qy	322	CCAAAGACCGTTTGTGTTAAGTTTGSTGTGAAATCTTTCCGCGTGAACACACAACTC	381	
Db	403	CCAAAAATGGGGTCTTCGCTCGCTGAGTAAATTTTCCGCGCTGACCTGGTCACTG	462	
Qy	382	CAAGAAGAACTCACAAAGTACTCTGTGCGCTGCAGGTGAGAGAGACTTGCTCAAGGC	441	
Db	463	CAAGAAGATGACAAAGGACCTGTGTTGCTTGCACAAAGAGAGACCTCTGGAAGAA	522	
Qy	442	AGTTGACGTTAATGACACCAAGCGAGCTCTTTGATTTACACATTGTGCAATGTAG	501	
Db	523	CGCTGACCTGCAAGGCGCAACACAGCGCCCTTCTCATATCCACTTCTGCAAGTGGAA	582	
Qy	502	ATTGGGATTTTGTATGAAGCCTTGGAGAGAGACCTTAGCAAAAAATTAATACATCT	561	
Db	583	ATCGGATTTATGATGAACCTCTGATCGAGAACACTCAAGACCAATGATATCTTGCC	642	
Qy	562	CAGCAAGACGCACTGAGAGACAAATGTGTGAATTTCAACATTAACACATTGACAAACA	621	
Db	643	AACCAAGGAAATCTTGAAGAAATATCTAGACTTCATCAGAGGCAACACGGGCAAGCT	702	
Qy	622	CCACAGATGAGATTTCCAGCTCTTGAAGATTTGSCCGTGGCTGAGATGATGAAATC	681	
Db	703	CCCGAGATGATTTCCAGGCTCTTGAATTTGAAAGAAAGCTGGAATATATGACATC	762	
Qy	682	CGTTTGACCCGCGCAAGAGCAGGGAAAGCAGAAATCAATCTGGCCGTTGCCAACAG	741	
Db	763	AGTTTTCACATGGCTTTCAGACAGAAAGGAGCAAAATTAATCTAGCAGTTTCTCACATG	822	
Qy	742	GGATTTTATGTTTTCAGGCTTTTACTTAAGATCAATGCTTCACTGGGCCAAGTGGCG	801	
Db	823	GGTGTCTCGGTGTTCCAGGGTACCAACAAATCAATATCTTCAACTGGTCCAAAGTCCGT	882	
Qy	802	AAGCTGAGCTTCAAGAGAAAGCGCTTCTCATCAAGCTCGGCGCAGATGCAATATGTCG	861	
Db	883	AAACTTATGCTTCAAGAGAAAGGTTTCTTATCAAACTCAACCTGAGAGTCCATAGGCGC	942	
Qy	862	TACCAAGATACCTTGAATTCCTGATGGCCAGTCCGGAATTTCTGCAGTCTTCTTGAAA	921	

Dh 943 TACCAGACAGTTAGATTCTTGTGGGTAGCAGAGTGAATGAAGAACTTCTGGAG 1002
Qy 922 ATCTGTGTGAACATCATGCTTCTTTAGACTTTTGAAGAGCCAAACCAAGCCAG 981
Db 1003 ATATGTGTGAATCACTACCTTTTAACTCTCTACAGCCTTAAGCCAAAGGCAAG 1062
Qy 982 CCGGTCTCTTTAGCCGGGGTCAATCATTTGCTTCAAGTGGTGGGACTCAGAAAGCAGTT 1041
Db 1063 GCTGTCTTCTTACGCCAGGCTCTCTTCAAGATACGTGAGAGAACTCAGAAACAACTA 1122
Qy 1042 CTGACTATGTAAAGAGAGACATAGAAGGTGAGTTGAAGAAAGCAAGCAGAG 1101
Db 1123 GTAGATTATGTCAAAAGACGGTGAATGAAGAAATTCATACGAAGACGGCAGATAG 1182
Qy 1102 ATTCT--ATTCTATCCGAGCTTGTCTTACAGCTTACAGAACTGAATTGGAAAGTCTG 1158
Db 1183 ACTCGACATCTCTTCAATGCTGACTGTAGATCTGCTTAAACAGACGCTCTCTTAC 1242
Qy 1159 GAGCAGTCTCAGACAGACAGCCTTAATTTGAGAGAGGTGCGAATCTCAGGGGGC 1218
Db 1243 GATGGCTTGAAGACTTGTGCTCTCTGTCTTCAAGAAATGTCTCTTTATTCAC----- 1297
Qy 1219 CAGAGCTGCCGGCAGAGAAAGAACGAAAGTTTCCGCCGGGAGCCGGGGTCCAGCCG 1278
Db 1298 -----CCCTAAGTTCTTTGTCTCTGCTGCTGCCCAATTGAAGGACAGCAGAG 1352
Qy 1279 AGCCCTGCCCGAGAGAAAGCCCGGGGTAAACAAGCGGAGCGAGCGGCGCTCGGCG 1338
Db 1353 CTCTCTGTGATCCGAGGCTCCCTCATCAAGAGCAGCAGCAGAGAGAGCAGCAG 1412
Qy 1339 CCCACGAGAGAAAGAGAGAGGTGTTAAGGATAGAGCCAGCAGAGTAAACCTCAGCCC 1398
Db 1413 ACCA--TCATCATCTGACGGGCCAGACACAGTGGCCCATCTCCCTGAGCCCCCGTG 1470
Qy 1399 CCGCAGCCAAAGCAGAGCTCCCTGACTGAGCTCCTCACTTTCGAGCTGTCTGTAGAC 1458
Db 1471 CTCCGGCTGTGATCAGGCTTTTCTATGAGATGTCTCAGGCTTCTCCCTCAGCTGAG 1530
Qy 1459 TCCGAGGGGGAGTGGCCCTGCGCAAGTGACTTGTCTCCCAACTGAGCCCGACACC 1518
Db 1531 AGCCACTGAGCCCTGTGCTGAGCTTCAAGCCCG-----CTGAGTACAGCTGAG 1581
Qy 1519 AAGCAGGCTCTCCCTTATGACCGGCTGTGAATGACAGCGCTCCCGGAGCGAGC 1578
Db 1582 CAGGTGCTATCCCTGTGCTGAGCCCTGTGCTGAGCTGAGCTGTGATCAGCCAGATGAC 1641
Qy 1579 GATGAGATGAGGCGCCGAGAGAGATTCCCACTGAATAAAGCTTCTATAGTAAAG 1638
Db 1642 AACCAAGAAAGAGGAAACACACATGCAATGAGATAGGCTATTTTCAATAGCCAG 1701
Qy 1639 GAAGTGTCTAACACCGAGCGAATATCTGAGAGATCTCGAATTAATCTTGTGTTT 1698
Db 1702 GAGATTCTCCCTACAGAAAGCACTATCTGAGATTTAGAAATTAACATGTTGTGTTT 1761
Qy 1699 CAGAGCAGAGGAGCAAGAGAGAGCGCAATGCCGAACCTGAAGAAATCTCATATTCCG 1758
Db 1762 CGAGGCTGCTGATCAAGAGAGAGAGCAATGCTGAGCCCTGATGAGCCCTGCTTTCTCC 1821
Qy 1759 AATTTGAACCTTGGACAAATTTTATATATTTTCTCAAGAAATTTGAGCAGACTT 1818
Db 1822 AACATGATCCAGTCTACAGATTCCACAGAGGCTTCTTCAAGAGTGAAGAGCTG 1881
Qy 1819 GCCCTGTGGAGAGCCGCTCAATATGCCA--AATCAGAGTTTACCAAGATTTGGGAT 1875
Db 1882 GCATCTGGGAGAGGCTCCAGATGCCCACTTAAAGATATCAACAGCAATGGGGAC 1941
Qy 1876 GTATGCTGAAGACATTTACAGGAGCATGAGCACTGGGCGCTCCTGTGGAGCAGAC 1935
Db 1942 ATCTCTCTCAGAGACATGCTGATTAAGAAATTTACTAGCTTCTTCAAGACAGAT 2001
Qy 1936 GAGGCTCTTGAAGGCTGTGAGAAATGAATCAAGAGCTCCCGCGGCTGAGAACTTCTGC 1995
Db 2002 GAGGTCTTAACAGAACTGGAAAGGCGCAAAACATCTGTAAAGAAAGCTGAGGAGCTTAC 2061

Qy 1996 AGAGACTTTGAGCTGAGAAAGGTGTGTACTTACCCGCTCAACACTTCTCTTGGGGCA 2055
Db 2062 AAAGATTTGAGCTCAAAAGGTGTGTGTACTTCTCTTCAACATTTCTGCTAAAGCC 2121
Qy 2056 CTGACCGGCTCATGACATCAAGACAGATCTTGAAGGAGCTGTGCAACACACCGCGC 2115
Db 2122 GTCCAGAGGCTATCCATTCAGCTGTGTGTGAGAGCGGCTGTGTGTGCTACTTCTCT 2181
Qy 2116 AGCCAGCGGACTTCAGGAGCTGCCAGCCGCTTTGGCAGAGATCAACGAGATGTGGCA 2175
Db 2182 GGGCAGCGGCACTATGCGCATGCTCATGAGGCACTGAAAGCCATCAAGAAATGACCA 2241
Qy 2176 CAGTCCACGCTACAGATGATCAAGATGAGAGAAATTTCCAGAAAGCTCAGAAA 2235
Db 2242 GAGCTCAGAAAGCTTACCCGCTGAGAAACCTTACAGAAATTTGACGAGCTACAGCA 2301
Qy 2236 GATTTGATTTGAGCTTGAATCTTGTGTGTTCCGGGAAAGGAGTTCAATCCGCTGGCAGC 2295
Db 2302 GACTGTGTGTGTGAAGAACTCATTTGCTCTGGAGGAGTTTATCCGTGAGGGCTGC 2361
Qy 2296 CTGACAGGCTCTCGGGGAAAGGGGCTTCAGACGCGATGTTCTTCTGTTCAGACGTC 2355
Db 2362 CTGACAGGCTCAACAAAGGGGCTGCAACAGAGATGTTTTTCTGTCTCAGATATG 2421
Qy 2356 CTGCTATACAGAGCCGGGGCTGACGGCTCCAATCAGTTTAAAGTCCAGGGGAGCTC 2415
Db 2422 TTGCTGTATCAAGCAAAAGTGTCAAGAGCCAGATATTTCCGAACTCTGTGCTCTT 2481
Qy 2416 CCGCTATGAGCTGACGATTTGAGAGAGCGAAGACGAGTGGGGGTGCCCACTGCTG 2475
Db 2482 CACTCCGTGTGCTGTGTATGAAGAAAGTGAAGAAATGATGTGTCTCTCATTTGCTTC 2541
Qy 2476 ACCCTCGGGGCAAGGGGCAATCATCTGTGAGCCGCACTTCTGTGTGAGATGAG 2535
Db 2542 ACCATCATGACACTCAAGAAACAAATGTGTGTGAGCCAGCACTCGGCTAGAAAGAA 2601
Qy 2536 AATGTGTTGAGACATTCAGATGCGCATTTGACCTGGCGGAGAAAGACAGACCCCGCC 2595
Db 2602 AAGTGAATGAGAGCCTGAATGAGCAATCAAGACCAAGCACTATCGGTACTACCC 2661
Qy 2596 CTTGAGTCTCTGCGCAGAGCCGCCCTGACAAAGTCCCTGATGAAAGCAGCGGGCT 2655
Db 2662 CAGTGTGTGTGGAGAGCCCGGTGTATCTCGTATCCCTGATCTTGTATGAAAGTCTCT 2721
Qy 2656 GACCAAGATCAGAGATGACTGAGCGCTGCGGCAATGCTGTGAGCGCCAGGCCCG 2715
Db 2722 CTGAAAGATCAGAAAGTGTGAGGAAAGCCGGGGCTGCTGTGAGGGGAAAGAGCCAG 2781
Qy 2716 CACCGCGG--CAACAAATGTGTCACTGTGTGCGGACCCGCAACACAGGCTCTCCATGG 2773
Db 2782 CACCGGGGCCCAATCAACATGATGTGTGTGTGTGCTGTATTAACAAGTGTGCCAGAG 2841
Qy 2774 TGAATTCAGCATGCACTGAGTGAAGATCAGTTGTCTGGAACCTGCTGAGAGAAATTCAAA 2833
Db 2842 CAAATCAGATGAGCTGTGTAAGAACAGCTTTCAGAGATATCTGCTGAGAAAGTTCAAG 2901
Qy 2834 AAGCAACGGGTGAGAGCTGTGGGTGTTCACAACTTCTGCTGTCTTCTTACA 2893
Db 2902 ACAATATGCTGGCAGAAAGCTCTGGGTGTCTTTCACAACTTCTGCTGTCTTCTATATA 2961
Qy 2894 AATCAACAGAGCAATCAATCCCTTGTGAGCGGCTGCTGCTGCTGCTGCTGCTGCA 2953
Db 2962 AAACACATTCAGATGACTACCCCTGCGAGCCCTCCCATATTTGGTTAAGAGCTGAGCC 3021
Qy 2954 TCCCTCTGAGTCCGAGAACATCCAGAAAGACTACAGTTCAGAGCTTCAAGTCCC 3013
Db 3022 TCCCGAGGAGGCTGACAGATTCACAAAGACTATGTCTTCAAGCTCAATTCATTC 3081
Qy 3014 ACGTCTATCTTCAAGGCGGAAAGAGTATCACTTCCAAAGGTGATGAAAGTATCC 3073
Db 3082 ACGTCTATCTTCCGGGCTGAGAGCAAGTACATTTGAAGGTGATGAGAGTCA 3141

QY 3074 GCAGTCCACACAGCTGTGCTCGGACCCACAGCTGTGAGCCACAAAGTCTCTT 3129
 DB 3142 AAGGCGCAGCACTCACCAGGAGACCCCAAGTTTCACTCAAGCTGTCACTCAT 3197
 RESULT 7
 LOCUS BX364606 968 bp mRNA linear EST 08-APR-2004
 DEFINITION BX364606 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
 CDNA clone CS0DC027YA10 5-PRIME, mRNA sequence.
 ACCESSION BX364606 GI:46289717
 VERSION BX364606.2
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homnidae; Homo.
 REFERENCE 1 (bases 1 to 968)
 Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 On May 5, 2003 this sequence version replaced gi:30370819.
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
 Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
 end enriched, double-strand cDNA was digested with Not I and cloned
 into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
 was normalized. Library was constructed by Life Technologies, a
 division of Invitrogen. This sequence belongs to sequence cluster
 7229.f

For more information about this cluster, see
 http://www.genoscope.cns.fr/cdnas=CS1AC007Z806P1&c=7229.f.
 Location/Qualifiers

FEATURES

source
 1..968
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DC027YA10"
 /tissue_type="NEUROBLASTOMA COT 25-NORMALIZED"
 /clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo(dT)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
 ORIGIN

Query Match 30.5%; Score 957; DB 4; Length 968;

Best Local Similarity 99.9%; Pred. No. 9,9e-242;

Matches 968; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1062 AGGACATTAAGAGGTGCGATTGAAAAGAAACACAGCAGATTCTTATCCGAGCCT 1121
 DB 1 AGGACATTAAGAGGTGCGATTGAAAAGAAACACAGCAGATTCTTATCCGAGCCT 60
 QY 1122 TGGTTACAGCCATACGAATCTCGAAGTCTGAGACAGTCTCAGAGAGCCACG 1181
 DB 61 TGGTTACAGCCATACGAATCTCGAAGTCTGAGACAGTCTCAGAGAGCCACG 120
 QY 1182 CCTTACATTTGAGAAAGTGCAGATCTCCAGGGGGCCAGAGTGCCTGGGAGAAAGA 1241
 DB 121 CCTTACATTTGAGAAAGTGCAGATCTCCAGGGGGCCAGAGTGCCTGGGAGAAAGA 180
 QY 1242 ACCGAAAGTTTCCGCGGGGAGCCGGGGGTGCACCCGAGCCTTGGCCGAGAAAGCCC 1301
 DB 181 ACCGAAAGTTTCCGCGGGGAGCCGGGGGTGCACCCGAGCCTTGGCCGAGAAAGCCC 239
 QY 1302 CGCGGGGTAAACAAGCGCGAGCGCGCTCGGGCCCAAGAGAAAGAGAGAGGT 1361
 DB 240 CGCGGGGTAAACAAGCGCGAGCGCGCTCGGGCCCAAGAGAAAGAGAGAGGT 299

QY 1362 CGTTAAGATTAAGACCCAGACAGATTAACCTCAGCCCCCGGACCGAAGCAGAGTCCCT 1421
 DB 300 CGTTAAGATTAAGACCCAGACAGATTAACCTCAGCCCCCGGACCGAAGCAGAGTCCCT 359
 QY 1422 GACTGGCAGTCTCTCACCTTTCCGAGCTGTCTGTAACTTGCAGGGGGAGGTGCCCCCTGC 1481
 DB 360 GACTGGCAGTCTCTCACCTTTCCGAGCTGTCTGTAACTTGCAGGGGGAGGTGCCCCCTGC 419
 QY 1482 CAACGAGACCTTGTCTCCCACTGAGCCCGGACACCAAGCAGGCTCTCCCTTGTACG 1541
 DB 420 CAACGAGACCTTGTCTCCCACTGAGCCCGGACACCAAGCAGGCTCTCCCTTGTACG 479
 QY 1542 CCCGCTGCTGAATGAACAGAGCCTGCCCCCGGACCGAAGATGAGATGAGGAGGAGAA 1601
 DB 480 CCCGCTGCTGAATGAACAGAGCCTGCCCCCGGACCGAAGATGAGATGAGGAGGAGAA 539
 QY 1602 GAGATTTCCCACTGATTAAGGCTACTTCACTAAGAAAGTGTCTACCAACGAGGAAAC 1661
 DB 540 GAGATTTCCCACTGATTAAGGCTACTTCACTAAGAAAGTGTCTACCAACGAGGAAAC 599
 QY 1662 ATATCTGAAGATCTCGAAGTTATCATCTCGTGGTTTCAGACACAGTGAAGAGAGGA 1721
 DB 600 ATATCTGAAGATCTCGAAGTTATCATCTCGTGGTTTCAGACACAGTGAAGAGAGGA 659
 QY 1722 CGCATGCGGAGAGCAGTGAAGAGTCTCATATTCGGAATTTTGAACCTTTGCACAAAT 1781
 DB 660 CGCATGCGGAGAGCAGTGAAGAGTCTCATATTCGGAATTTTGAACCTTTGCACAAAT 719
 QY 1782 TCATCTAATTTTTCAGAGAAATTGAGCAAGACTTGGCCCTGTGGAGAGGCCGCTCAAA 1841
 DB 720 TCATCTAATTTTTCAGAGAAATTGAGCAAGACTTGGCCCTGTGGAGAGGCCGCTCAAA 779
 QY 1842 TGGCCCAATCAGAGATTACCAAGAATCGGCGATGCTGTAAGAAATTCAGGGCAT 1901
 DB 780 TGGCCCAATCAGAGATTACCAAGAATCGGCGATGCTGTAAGAAATTCAGGGCAT 839
 QY 1902 GAGCAGCTGGGCGGCTCATCTGTGAAGACAGCGAGCCTTGAAGGCCCTTGAGAGATG 1961
 DB 840 GAGCAGCTGGGCGGCTCATCTGTGAAGACAGCGAGCCTTGAAGGCCCTTGAGAGATG 899
 QY 1962 AATCAAGAGCTCCCGGCGGCTGTGAAGAACTTCTGCAAGACTTTGAGCTGCAAGGTGG 2021
 DB 900 AATCAAGAGCTCCCGGCGGCTGTGAAGAACTTCTGCAAGACTTTGAGCTGCAAGGTGG 959
 QY 2022 TTACCTACC 2030
 DB 960 TTACCTACC 968

RESULT 8

BX354921

LOCUS

BX354921 1000 bp mRNA linear EST 23-APR-2004

DEFINITION BX354921 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens

CDNA clone CS0DC027YA10 5-PRIME, mRNA sequence.

ACCESSION BX354921

VERSION BX354921.2 GI:46553760

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Homnidae; Homo.

REFERENCE 1 (bases 1 to 1000)

Li, W.B., Gruber, C., Jessee, J., and Polayes, D.

Full-length cDNA libraries and normalization

Unpublished (2001)

On May 5, 2003 this sequence version replaced gi:30383921.

COMMENT Contact: Genoscope

Genoscope - Centre National de Sequencage

2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE

Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr

1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime

end enriched, double-strand cDNA was digested with Not I and cloned

into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 7229.f

For more information about this cluster, see <http://www.genoscope.cns.fr/cdnats=CS0PC027BA05OP1&c=7229.f>.

FEATURES

source

```

1. 1000
   /organism="Homo sapiens"
   /mol_type="mRNA"
   /db_xref="taxon:9606"
   /clone="CS0DC027YA10"
   /issue_type="NEUROBLASTOMA COT 25-NORMALIZED"
   /clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
   /note="Left strand cDNA was primed with a NotI-oligo (dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

```

ORIGIN

```

Query Match      30.0%; Score 940.2; DB 4; Length 1000;
Best Local Similarity 99.2%; Pred. No. 2.9e-237;
Matches 961; Conservative 4; Mismatches 2; Indels 2; Gaps 2;

1062 AGGACATPAAGAGGTGCGAGTTTGAAGAGAGCAGCAGAAATTCATTCTATCCGAGCCT 1121
    |||
    1 AGGACATPAAGAGGTGCGAGTTTGAAGAGAGCAGCAGAAATTCATTCTATCCGAGCCT 60
    |||
    1122 TGGCTTCACAGCCTTACAGAACTGAATTCGGAAGTCTGAGAGCTCTCAGCAGACACAG 1181
    |||
    61 TGGCTTCACAGCCTTACAGAACTGAATTCGGAAGTCTGAGAGCTCTCAGCAGACACAG 120
    |||
    1182 CCTTACCTTTGGAGAGAGTCCGGAATCTCCAGGGGGGCGAGAGCTGCGGGGAGAGAAAGA 1241
    |||
    121 CCTTACCTTTGGAGAGAGTCCGGAATCTCCAGGGGGGCGAGAGCTGCGGGGAGAGAAAGA 180
    |||
    1242 ACCGAAGTTTCCGCGGGGAGCGGGGAGTGCACCCGAGCCCTGCGCGGAGAGAAAGCCC 1301
    |||
    181 ACCGAAGTTTCCGCGGGGAGCGGGGAGTGCACCCGAGCCCTGCGCGGAGAGAAAGCCC 239
    |||
    1302 CCGCGGTTAACAGAGCGGAGCGGAGCGGCTCGGCGCCGACGAGAGAGAGAGAGAGT 1361
    |||
    240 CCGCGGTTAACAGAGCGGAGCGGAGCGGCTCGGCGCCGACGAGAGAGAGAGAGAGT 299
    |||
    1362 CGTTAAGATAGAGACCCAGCAGAGATAAAGTCAAGCCCCCGGACGACAGAGCTCCCT 1421
    |||
    300 CGTTAAGATAGAGACCCAGCAGAGATAAAGTCAAGCCCCCGGACGACAGAGCTCCCT 359
    |||
    1422 GACTGGAGTCTCAACCTTCCGAGCTGTCTGTAATTCGAGAGGGGAGAGGCGCCCTGC 1481
    |||
    360 GACTGGAGTCTCAACCTTCCGAGCTGTCTGTAATTCGAGAGGGGAGAGGCGCCCTGC 419
    |||
    1482 CAAAGTGAAGCTTGTCTCCCAAGCTGAGCCCGGACACCAAGAGGCTCTCCCTTGATCAG 1541
    |||
    420 CAAAGTGAAGCTTGTCTCCCAAGCTGAGCCCGGACACCAAGAGGCTCTCCCTTGATCAG 479
    |||
    1542 CCGCGTCTGAATGACAGGCTGCGCGGAGCGGAGCAGATGAGAGATAGGGCCGAGAGAA 1601
    |||
    480 CCGCGTCTGAATGACAGGCTGCGCGGAGCGGAGCAGATGAGAGATAGGGCCGAGAGAA 539
    |||
    1602 GAGATTCGCAAGTATGAAGGATCTTATAGCTTAAGAGAGTGTCTACACCGAGCGAGAC 1661
    |||
    600 ATATCTGAAGATCTCGAAGTATCACTTGTGTGTTTCAAGACCACTGAGCAAAAGAGA 1721
    |||
    1722 CGGCAATCCGGAAGCACTGAAAGTCTCATATTCGCAATTTTGAACCTTTGCACAAAT 1781
    |||
    660 CGGCAATCCGGAAGCACTGAAAGTCTCATATTCGCAATTTTGAACCTTTGCACAAAT 719
    |||
    1782 TCATTAATTAATTTCTCAAGAAATTAAGACGACTTGCCCTGTGGAGAGGCGGCTCAA 1841

```

```

Db      720 TCATTAATTAATTTCTCAAGAAATTAAGACGACTTGCCCTGTGGAGAGGCGGCTCAA 779
    |||
    1842 TGGCCAAATCAGAGATTACCAAGAAATCGCGGATGATGATCTGAAGAACTTCAGGGCAT 1901
    |||
    780 TGGCCAAATCAGAGATTACCAAGAAATCGCGGATGATGATCTGAAGAACTTCAGGGCAT 839
    |||
    1902 GAAGCACTGCGGGCTCAGCTGTGGAAGCAGCAGAGGCGC-TTGGAGGCGCTGGAGAAATG 1960
    |||
    840 GAAGCACTGCGGGCTCAGCTGTGGAAGCAGCAGAGGCGCTTGGAGGCGCTGGAGAAATG 899
    |||
    1961 GAATCAAGAGCTCCGCGCGCTGAGAACTTCTGCAAGAGACTTGAAGTCTGAGAGGCTGT 2020
    |||
    900 GAATCAAGAGCTCCGCGCGCTGAGAACTTCTGCAAGAGACTTGAAGTCTGAGAGGCTGT 959
    |||
    2021 GTTACTTAC 2029
    |||
    960 GTTACTTAC 968

```

```

RESULT 9
B0881649
LOCUS
DEFINITION
ACCESSION
VERSION
B0881649.1 GI:22273657
KEYWORDS
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
1 (bases 1 to 912)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LIML at:
http://image.llnl.gov
Plate: L1CM2533 row: n column: 12
High quality sequence stop: 684.
Location/Qualifiers
1. 912
   /organism="Homo sapiens"
   /mol_type="mRNA"
   /db_xref="taxon:9606"
   /clone="IMAGE:6339275"
   /issue_type="neuroblastoma, cell line"
   /lab_host="DH10B (phage-resistant)"
   /clone_lib="NIH_MGC_47"
   /note="Organ: Brain; Vector: pOTB7; Site_1: XhoI; Site_2:
   EcoRI; cDNA made by oligo-dT priming. Directionally
   cloned into EcoRI/XhoI sites using the following 5'
   adaptor: GGACGAG(G). Size-selected >500bp for average
   insert size 1.8kb. Library constructed by Ling Hong in
   the laboratory of Gerald M. Rubin (University of
   California, Berkeley) using ZAP-cDNA synthesis kit
   (Stratagene) and Superscript II RT (Life Technologies).
   Note: this is a NIH_MGC Library."

```

```

ORIGIN
Query Match      28.2%; Score 883.4; DB 3; Length 912;
Best Local Similarity 98.7%; Pred. No. 3.2e-222;
Matches 901; Conservative 0; Mismatches 11; Indels 1; Gaps 1;

544 AAAATTAATATACCTCTCAAGCAAGCAGCACTAGAGACAAATCTGTGAATTTACCAT 603

```

```
Db      1 AAAAAAAAAATACCTACTGCAAGACGCACTAGAGACAAATTCGTGAATTTCAACAT 60
Qy      604 AACCACTTGGACAAACACCAAGCAATGATTTTCCAGTCTTAAGATATGCCCTCGG 663
Db      61 AACCAATTTGGACAAACACCAAGCAATGATTTTCCAGTCTTAAGATATGCCCTCGG 120
Qy      664 CTAGAGATGATGAAATCCGGTTGCAACCCGGCCAAAGACAGGAGGACAGATCAAT 723
Db      121 CTAGAGATGATGAAATCCGGTTGCAACCCGGCCAAAGACAGGAGGACAGATCAAT 180
Qy      724 CTGGCCGTTGCCAAACCGGAAATTTCTAGTGTTCAGGGTTTCAAGATCAATGCCCTTC 783
Db      181 CTGGCCGTTGCCAAACCGGAAATTTCTAGTGTTCAGGGTTTCAAGATCAATGCCCTTC 240
Qy      784 AACTGGGACCAAGTGGCGGAAGCTGAGCTTCAAGAGAAAGGCTTTCTCAATCAAGCTCCCG 843
Db      241 AACTGGGACCAAGTGGCGGAAGCTGAGCTTCAAGAGAAAGGCTTTCTCAATCAAGCTCCCG 300
Qy      844 CCAGATGCCAATAGTGGTACCAAGATACCTTGAATTCCTGATGGCCAGTCCGGATTTTC 903
Db      301 CCAGATGCCAATAGTGGTACCAAGATACCTTGAATTCCTGATGGCCAGTCCGGATTTTC 360
Qy      904 TGCAGTCTCTTGTGAAAATCTGTGTGAACATCATGCTCTTTAGACTTTTGAAGAG 963
Db      361 TGCAGTCTCTTGTGAAAATCTGTGTGAACATCATGCTCTTTAGACTTTTGAAGAG 420
Qy      964 CCGAAACCAAAAGCCCAAGCCCGTCTTTAGCCGGGGGTATCATTTGCGTTCAATGCT 1023
Db      421 CCGAAACCAAAAGCCCAAGCCCGTCTTTAGCCGGGGGTATCATTTGCGTTCAATGCT 480
Qy      1024 CGGACTCAGAGAGAGTTCTGCACTATGTTAAAGAGAGACATTAAGAGAGTGCAGTTT 1083
Db      481 CGGACTCAGAGAGAGTTCTGCACTATGTTAAAGAGAGACATTAAGAGAGTGCAGTTT 540
Qy      1084 GAAAGAGACACAGCAAGATTCAATTCTATCCGAGCTTGTCTTCAAGCTTACAGAACTG 1143
Db      541 GAAAGAGACACAGCAAGATTCAATTCTATCCGAGCTTGTCTTCAAGCTTACAGAACTG 600
Qy      1144 AATTGCGAAGTGTGAGAGAGTCTCAGACAGACACCAAGCTTTACATTTGAGAGAGTGGC 1203
Db      601 AATTGCGAAGTGTGAGAGAGTCTCAGACAGACACCAAGCTTTACATTTGAGAGAGTGGC 660
Qy      1204 GAATCTTCAGAGGGGACGAGAGCTCCGGCGAGAGAAAGAACCGAAGTTTCCGCGGGAG 1263
Db      661 GAATCTTCAGAGGGGACGAGAGCTCCGGCGAGAGAAAGAACCGAAGTTTCCGCGGGAG 720
Qy      1264 CCGGGGTGCGACCCGAGCCCTCGCGCGAGAGAAAGCCCGGGGTAAACAAGCAGCGGAGC 1323
Db      721 CCGGGGTGCGACCCGAGCCCTCGCGCGAGAGAAAGCCCGGGGTAAACAAGCAGCGGAGC 780
Qy      1324 GAGAGCGGCTCGGCGCCCAAGAGAGAAAGAGAGAGTGTGTTAAGATAGAGACCAAGCAG 1383
Db      781 GAGAGCGGCTCTCGGCGCCCAAGAGAGAAAGAGAGAGTGTGTTAAGATAGAGACCAAGCAG 840
Qy      1384 AGTAAACCTTCAAGCCCGCAGCAAGCAAGCAAGCTCCCTGACTGCAAGTCTCTCTTCC 1443
Db      841 AGTAAACCTTCAAGCCCGCAGCAAGCAAGCAAGCTCCCTGACTGCAAGTCTCTCTTTC 899
Qy      1444 GAGCTGTCTGTGA 1456
Db      900 GAGCTGTCTGTGA 912

RESULT 10
BUI31960      904 bp      mRNA      linear      EST 18-OCT-2002
LOCUS      AGENCOURT 10485886 NIH MGC 109 Homo sapiens cDNA clone
DEFINITION      IMAGE:6672254 5', mRNA sequence.
ACCESSION      BUI31960
VERSION      BUI31960.1 GI:24120779
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
```

```
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo
REFERENCE      1 (bases 1 to 904)
AUTHORS      NIH-MGC http://mgc.nci.nih.gov/.
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL      Unpublished (1999)
COMMENT      Contact: Robert Strausberg, Ph.D.
              Email: cgapbs-remail.nih.gov
              Tissue Procurement: ATCC
              cDNA Library Preparation: Rubin Laboratory
              cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
              DNA Sequencing by: Agencourt Bioscience Corporation
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LNL at:
              http://image.llnl.gov
              Plate: LNCM2952 row: h column: 14
              High quality sequence stop: 721.
              Location/Qualifiers
                1..904
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="IMAGE:6672254"
                /tissue_type="teratocarcinoma, cell line"
                /lab_host="DH10B (phage-resistant)"
                /clone_1ib="NIH MGC 109"
                /note="Organ: ovary; Vector: pORF7; Site_1: EcoRI; Site_2:
                XhoI; cDNA made by oligo-dT priming. Directionally cloned
                into EcoRI/XhoI sites using the following 5' adaptor:
                GGCAAGAG(G). Library constructed by Ling Hong in the
                laboratory of Gerald M. Rubin (University of California,
                Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
                Superscript II RT (Life Technologies). Note: this is a
                NIH-MGC library."

ORIGIN
Query Match      27.9%; Score 873.2; DB 3; Length 904;
Best Local Similarity 98.6%; Pred. No. 1.6e-219;
Matches 891; Conservative 0; Mismatches 10; Indels 3; Gaps 1;

Qy      967 AAACCAAGCCCAAGCCGCTCTTTAGCCGGGGGTATCATTTGCGTTCAAGTGTCCG 1026
Db      1 AAACCAAGCCCAAGCCGCTCTTTAGCCGGGGGTATCATTTGCGTTCAAGTGTCCG 60
Qy      1027 ACTCAGAACAGATTCTGACATATGTTAAAGAGAGACATTAAGAGAGTGCAGTTTGA 1086
Db      61 ACTCAGAACAGATTCTGACATATGTTAAAGAGAGACATTAAGAGAGTGCAGTTTGA 120
Qy      1087 AGAAGACACAGCAAGATTCAATTCGAGAGCTTGTCTTCAAGCTTACAGAACTGAAT 1146
Db      121 AGAAGACACAGCAAGATTCAATTCGAGAGCTTGTCTTCAAGCTTACAGAACTGAAT 180
Qy      1147 TCGGAAAGTCTGAGAGCTTCAAGACAGACCAAGCTTTCAATTGAGAGAGTGCAGAA 1206
Db      121 AGAAGACACAGCAAGATTCAATTCGAGAGCTTGTCTTCAAGCTTACAGAACTGAAT 180
Qy      1207 TCTTCAGGGGGGACAGAGCTGCGCGAGAGAAAGAAACCGAAGTTTCCGCGGGAGCCG 1266
Db      181 TCGAAGTCTGAGAGAGTCTCAGACAGACACCAAGCTTTCAATTGAGAGAGTGCAGAA 240
Qy      241 TCTTCAGGGGGGACAGAGCTGCGCGAGAGAAAGAAACCGAAGTTTCCGCGGGAGCCG 300
Db      1267 GGGTCGACCCGAGCCCTGCGCGAGAGAAAGCCCGCGGGTAAACAAGCAGCGACGGA 1326
Qy      301 GGGTCGACCCGAGCCCTGCGCGAGAGAAAGCCCGCGGGTAAACAAGCAGCGACGGA 360
Db      1327 GCCGCTCGGCGCCCAAGAGAGAAAGAGAGTGTGTTAAGATAGAACCCAGCAAGT 1386
Qy      361 GCCGCTCGGCGCCCAAGAGAGAAAGAGAGTGTGTTAAGATAGAACCCAGCAAGT 420
Db      1387 AAACCTCAGCCCCCGAGCAAGCAAGCAAGCTCCCTGACTGCGAGTCTCACTTTCGAG 1446
Qy      421 AAACCTCAGCCCCCGAGCAAGCAAGCAAGCTCCCTGACTGCGAGTCTCACTTTCGAG 480
```


1447 CTGCTGGAAGTGGAGGGGGAGTGGCCCTGCGCAACGTCCTTGTCTCCAACTG 1506
Db CTGCTGGAAGTGGAGGGGGAGTGGCCCTGCGCAACGTCCTTGTCTCCAACTG 540
Qy 1507 AGCCCCGACACCAAGCAGGCTCTCCCTTGTATGAGCCCGCTGTGAATGACGAGGCTGC 1566
Db 541 AGCCCCGACACCAAGCAGGCTCTCCCTTGTATGAGCCCGCTGTGAATGACGAGGCTGC 600
Qy 1567 CCCCCGACGAGATGAGAGTGGGCGCGGAGAAAGATTTCCCACTGATTAAGCTGAC 1626
Db 601 CCCCCGACGAGATGAGAGTGGGCGCGGAGAAAGATTTCCCACTGATTAAGCTGAC 660
Qy 1627 TTCTAGCTAGAGAGTGTCTTACCAACGAGGAACTATCTGAAGATCTGAAGTTATC 1686
Db 661 TTCTAGCTAGAGAGTGTCTTACCAACGAGGAACTATCTGAAGATCTGAAGTTATC 720
Qy 1687 ACTTCGTGTTTCAGAGCAGTGAAGCAAGAGAGCGCATGCGGAAAGCTGAAAAAGT 1746
Db 721 ACTTCGTGTTTCAGAGCAGTGAAGCAAGAGAGCGCATGCGGAAAGCTGAAAAAGT 780
Qy 1747 CTCATATTCGGAATTTTGAACCTTTTGCAAAATTTCTACTAATTTTCTCAGAAATT 1806
Db 781 CTCATATTCGGAATTTTGAACCTTTTGCAAAATTTCTACTAATTTTCTCAGAAATT 840
Qy 1807 GAGCAAGACTTGGCC--TGTGGAAAGCGCGCTCAATGCCCAATCAGAGATTACCA 1863
Db 841 GAGCAAGACTTGGCCCTGCTGGAGAGGCGCGCTCAATGCCCAATCAGAGATTACCA 900
Qy 1864 AGAA 1867
Db 901 NGAA 904

RESULT 11
BM51766 1026 bp mRNA linear EST 20-FEB-2002
LOCUS AGENCOURT_6575294 NIH_MGC_98 Homo sapiens cDNA clone IMAGE:5479350
DEFINITION 5', mRNA sequence.
ACCESSION BM51766
VERSION BM51766.1 GI:18789071
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 1026)
NIH-MGC http://mgc.ncl.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cga@p-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LCM2000 row: h column: 07
High quality sequence stop: 675.
Location/Qualifiers

FEATURES
Source
1..1026
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5479350"
/tissue_type="astrocytoma grade IV, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_id="NIH_MGC_98"
/note="Organ: brain; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally

cloned into EcoRI/XhoI sites using the following 5'
adapter: GGACGAG(G). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC library."

Query Match 27.6%; Score 863.8; DB 2; Length 1026;
Best Local Similarity 96.7%; Pred. No. 5.1e-217;
Matches 916; Conservative 0; Mismatches 22; Indels 9; Gaps 3;
ORIGIN
Qy 1997 GAGACTTGGAGCTCAGAGAGTGTGTACCTACCGCTCAACACTTCTCTGCGGCAC 2056
Db 1 GAGACTTGGAGCTCAGAGAGTGTGTGTACCTACCGCTCAACACTTCTCTGCGGCAC 60
Qy 2057 TGACCGGCTCATATGCACTTACAGAGAGGCTCTGAGCGGCTGTGCAACACCAACCGCGGA 2116
Db 61 TGACCGGCTCATATGCACTTACAGAGAGGCTCTGAGCGGCTGTGCAACACCAACCGCGGA 120
Qy 2117 GCCACGCCGACTTACAGGAGTCTCCGAGCGCTTTGGAGAGATACAGAGATGTGTGAC 2176
Db 121 GCCACGCCGACTTACAGGAGTCTCCGAGCGCTTTGGAGAGATACAGAGATGTGTGAC 180
Qy 2177 AGCTCCAGGATACATGATGATCAAGATGAGAAATTTCCAGAGCTSCAGAACTCAAGAAAG 2236
Db 181 AGCTCCAGGATACATGATGATCAAGATGAGAAATTTCCAGAGCTSCAGAACTCAAGAAAG 240
Qy 2237 ATTTGATTTGCAATTTGACAAATCTTGTGTTCCGGAGAGGAGATTATCGTCTGGGCAAGC 2296
Db 241 ATTTGATTTGCAATTTGACAAATCTTGTGTTCCGGAGAGGAGATTATCGTCTGGGCAAGC 300
Qy 2297 TCAGCAAGCTCTCGGGGAGAGGGGCTCAGAGAGCGCATGTTCTTCTCTTCAAGACGTCC 2356
Db 301 TCAGCAAGCTCTCGGGGAGAGGGGCTCAGAGAGCGCATGTTCTTCTCTTCAAGACGTCC 360
Qy 2357 TGCTATACAGAGAGCGGGGCTGAGCGGCTCAATGATTAAAGTCCAGCGGAGGCTCC 2416
Db 361 TGCTATACAGAGAGCGGGGCTGAGCGGCTCAATGATTAAAGTCCAGCGGAGGCTCC 420
Qy 2417 CGCTATATGCAATGACATTTGAGAGAGCGAAGCGAGTGGGGGTCGCCACTGCTGA 2476
Db 421 CGCTATATGCAATGACATTTGAGAGAGCGAAGCGAGTGGGGGTCGCCACTGCTGA 480
Qy 2477 CCTCCGGGGCCAGCGGCACTCATCTGTCGCGCCGCACTTCTCGGATCGAATGAGAG 2536
Db 481 CCTCCGGGGCCAGCGGCACTCATCTGTCGCGCCGCACTTCTCGGATCGAATGAGAG 540
Qy 2537 AGTGGGTTGAGGACATCCAGATGGCCATTGACCTGGGCGAGAGAGAGAGAGCGCCGCC 2596
Db 541 AGTGGGTTGAGGACATCCAGATGGCCATTGACCTGGGCGAGAGAGAGAGAGAGCGCCGCC 600
Qy 2597 CTGAGTTCTTGCGCAGCAGGCCCCCTTGACAACAAGTCCCTGTATGAAGCAGCGGCTG 2656
Db 601 CTGAGTTCTTGCGCAGCAGGCCCCCTTGACAACAAGTCCCTGTATGAAGCAGCGGCTG 660
Qy 2657 ACCAGAGTCAAGAGATGACTGAGCGCTCGGCGCAGATGCTGTGAGAGCGGCGCCGCC 2716
Db 661 ACCAGAGTCAAGAGATGACTGAGCGCTCGGCGCAGATGCTGTGAGAGCGGCGCCGCC 720
Qy 2717 ACCGCGGCAACAATATGTCAGTGTGTCGAGACCGCAACACAGCGTCTCCATGCTGG 2776
Db 721 ACCGCGGCAACAATATGTCAGTGTGTCGAGACCGCAACACAGCGTCTCCATGCTGG 780
Qy 2777 ACTTCAGCATCGAGTGAAGATCACTTGTCTGAAAACCTGTGAGGAAATTCAAAAACA 2836
Db 781 ACTTCAGCATCGAGTGAAGATCACTTGTCTGAAAACCTGTGAGGAAATTCAAAAACA 840
Qy 2837 GCACAC--GGGTGGCAAGCTGTGTGTGTTCACAAACTTTCGCTG-----TTCTTCT 2890
Db 841 GCACACGGGTGGCAAGCTGTGTGTGTTCACAAAACTTTCGCTGTTTCTTCT 900
Qy 2891 ACAATACACACGAGCAAT---CATCCCTTCCAGCTGCTCTG 2934

Db 901 ACAATCACACGAGGACATTTCCTCCCTGGCCAGCCTGCTCTG 947

RESULT 12
BX354920/c
LOCUS
DEFINITION
BX354920 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
CDNA clone CS0DC027YA10 3-PRIME, mRNA sequence.
ACCESSION
BX354920 GI:46552853
VERSION
BX354920.2 GI:46552853
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.
1 (bases 1 to 1072)
Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On May 5, 2003 this sequence version replaced gi:30381927.
COMMENT
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
7229.f
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?c=CS0DC027BA0SNP1&c=7229.f.
FEATURES
source
Location/Qualifiers
1..1072
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DC027YA10"
/issue_type="NEUROBLASTOMA COT 25-NORMALIZED"
/clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
Query Match 27.2%; Score 852.4; DB 4; Length 1072;
Best Local Similarity 98.6%; Pred. No. 5.5e-214;
Matches 897; Conservative 4; Mismatches 5; Indels 4; Gaps 4;

QY 2226 ACTCAAGAAAGATTGATTGGCATGCAATCTTGTGCTCCGGAAAGGAGTTTCATCCG 2285
Db 1002 ACTCAARAAAAGATTAAAT-GCATTTACAACTTGTGGTT-CGGGAAAGAGTTTCATCCG 945

QY 2286 TCTGGGAGGCTCAGCAAGCTCTCGGGGAGGGGCTCAGACAGGAGTGTCTTCCGTT 2245
Db 944 TCTGGGAGGCTCAGCAAGCTCTCGGGGAA-GGGCTCAGACAGGAGTGTCTTCCGTT 886

QY 2346 CAACGAGCTCTCTATACACGAGCCGGGGGCTGACGGCTCCATCACTGTTAAAGTCCA 2405
Db 885 CAACGAGCTCTCTATACACGAGCCGGGGGCTGACGGCTCCATCACTGTTAAAGTCCA 826

QY 2406 CGGGCAGCTCCCGCTCTATGCGATGACGATTGAGAGAGGAAAGAGATGGGGGGTCC 2465
Db 825 CGGGCAGCTCCCGCTCTATGCGATGAGATGAGAGAGGAAAGAGATGGGGGGTCC 767

QY 2466 CCACTGCTGACCTCCGGGGCCAGGGGAGTCCATCATGTTGGCCGCGAGTTCTCGGTC 2525
Db 766 CCACTGCTGACCTCCGGGGCCAGGGGAGTCCATCATGTTGGCCGCGAGTTCTCGGTC 707

QY 2526 CGAGATGAGAAAGTGCGTTGAGGACATCCAGATGGCCATTGACCTGGCGGAGAGACAG 2585

Db 706 CGAGATGAGAAAGTGCGTTGAGGACATCCAGATGGCAATTGACCTGGCGGAGAGACAG 647

QY 2586 CAGCCCCGCCCTGAGATTCTTGCCAGCGAGCCCCCTGTGCAACAAGTCCCCCTGATGAAC 2645
Db 646 CAGCCCCGCCCTGAGATTCTTGCCAGCGAGCCCCCTGTGCAACAAGTCCCCCTGATGAAC 587

QY 2646 CACCGGCGGTGACAGAGGATGAGAGGATGACCTGAGCGGCTCGCGGACATCGCTGAGCG 2705
Db 586 CACCGGCGGTGACAGAGGATGAGAGGATGACCTGAGCGGCTCGCGGACATCGCTGAGCG 527

QY 2706 CCAAGCCCCGCAACCGGGCAACAACAATGGTGACGTGTGCTGGACCGCAACACGCGT 2765
Db 526 CCAAGCCCCGCAACCGGGCAACAACAATGGTGACGTGTGCTGGACCGCAACACGCGT 467

QY 2766 CTCATGCTGACCTTCAGATGCGAGTGGAGAAATCGTTGTGGAACCTGTGAGGA 2825
Db 466 CTCATGCTGACCTTCAGATGCGAGTGGAGAAATCGTTGTGGAACCTGTGAGGA 407

QY 2826 ATTCAAAAACAGCAACGGGTGGCAGAAAGCTGTGGGTGTTCACAAACTTGTGCTGTT 2885
Db 406 ATTCAAAAACAGCAACGGGTGGCAGAAAGCTGTGGGTGTTCACAAACTTGTGCTGTT 347

QY 2886 CTTTCAAAATACACACAGCAACATCATCCCTTGGACCTGCTTGTGCTGCTACTC 2945
Db 346 CTTTCAAAATACACACAGCAACATCATCCCTTGGACCTGCTTGTGCTGCTACTC 287

QY 2946 GCTCCACATCCCTCTGATTCGGAACATCCAGAAAGCTACGTGTTAAGCTGACCT 3005
Db 286 GCTCCACATCCCTCTGATTCGGAACATCCAGAAAGCTACGTGTTAAGCTGACCT 227

QY 3006 CAAATCCAGCTACTACTTACAGGGCGGAAAGCGAGTACAGTTGAAAGGTGATGA 3065
Db 226 CAAATCCAGCTACTACTTACAGGGCGGAAAGCGAGTACAGTTGAAAGGTGATGA 167

QY 3066 AGTATCCGCAAGTCCACAGCTCTGCTCGGACCCACGTTTGAGCCCAAGAGTC 3125
Db 166 AGTATCCGCAAGTCCACAGCTCTGCTCGGACCCACGTTTGAGCCCAAGAGTC 107

QY 3126 TCTTGTGTAAT 3135
Db 106 TCTTGTGTAAT 97

RESULT 13
AY415515
LOCUS
DEFINITION
Homo sapiens FAR2 gene, VIRUTAL TRANSCRIPT, partial sequence.
ACCESSION
AY415515
VERSION
AY415515.1 GI:39771474
KEYWORDS
GSS.
SOURCE
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.
1 (bases 1 to 3165)
Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,
Todd, M.A., Tanenbaum, D.M., Civallo, D.R., Lu, F., Murphy, B.,
Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Shtinsky, J.J.,
Adams, M.D. and Cargill, M.
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios

TITLE
JOURNAL
PUBMED
REFERENCE
AUTHORS
TODD, M.A., TANENBAUM, D.M., CIVELLO, D.R., LU, F., MURPHY, B.,
FERREIRA, S., WANG, G., ZHENG, X.H., WHITE, T.J., SHINSKY, J.J.,
ADAMS, M.D. and CARGILL, M.
TITLE
JOURNAL
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,

Rockville, MD 20850, USA
This sequence was made by sequencing genomic exons and ordering
them based on alignment.

FEATURES
Source
1. .3165
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1..>3165
/gene="PARP2"
/locus_tag="HCM5583"

ORIGIN

Query Match 26.9%; Score 842.4; DB 14; Length 3165;
Best Local Similarity 51.9%; Pred. No. 3.3e-211;
Matches 1606; Conservative 0; Mismatches 1444; Indels 45; Gaps 7;

Qy 21 GCCGACCCCGAGATCAAGACTGGGGGCCCGGAAATTCGGGATCAGTACCTTGAACG 80
Db 33 GCAGACTGCAAGGATGGCTTGGGTGCCAGACCCCTGTGGAGTTAGCACCTTGAAGC 92

Qy 81 TGGACAGAGCCGCCCAACCTTCAAGAAACTGGTCCATCAAAATCCAGATGCT 140
Db 93 TGGGCAAGACTCTTGGCCAGAAATGCAAGAGACCTGCACCTCAGAGTAAAGCTGCT 152

Qy 141 GGAATGACCCAGGAGGATTTGAAGTTCCAAAGAGCTCTGGGAAAGGTCGCTGGA 200
Db 153 GGAACACACATGGAATATTTGACATTGAGCTTAAATGCGATGCCAGGATTAATGAC 212

Qy 201 TGCAGTTGGCAACCACTCAACCTCGTGAAGTGAATTTGGCTCGAGTTCTCTGA 260
Db 213 ACAAGTGGAAAGGTTTAAACCTGTGAATGTGACTACTGGGATGAGTTTCAAAA 272

Qy 261 TCACAAAAGATCAAGGTGGTGGTGAATCTCTTAAACCAATGTGAACAGATTAG 320
Db 273 TACTCACTCTACTGGAATTTGGCTGGAACCTATGAAACCATCATTTAGGCAATACG 332

Qy 321 GCCAAGCAGGTTGTGTTAAGTTGTGTTGAATTTCTTCGCGTGAACACACAGACT 380
Db 333 GCCAAGAAATGTGTGCTTGCCTAGCTGTAAATTTTTCACCTATCTGTGTGACT 392

Qy 381 CCAAGAAAGATCAAGAGTACTGTTCGCGCTGCAGGTGAAGCAGAGACTTGGCTCA 440
Db 393 ACAAGAAAGATATACAGATATCTTGTTCCTTGAATTAAGAGAACCTGTGGAAGA 452

Qy 441 CAGGTGAGCTGTATGACACACAGCGAGCTCTTGAATTTACACATTTGTGCAAT 500
Db 453 GCGTTTGACCTGTGTGACACACAGCGGCTTCTCAGTCCCATCTCTGCAGTGGA 512

Qy 501 GATTGGGATTTTGAAGCTTGTGAAGCTTGAAGCAGAGCACTTGAAGAAAATAAT 560
Db 513 AATAGAGATTAAGATGAACGCTGAGCCAGAGCACTCAAGTGAACGATTTTGGC 572

Qy 561 TCAGCAAGACGCACTAGAGCAAAATCGTGAATTTCAACATTAACCAATTGGACA 620
Db 573 TGGCCAGCAGACATCGCTTGAAGAGATAGTAATTCATCAAGAGACATGGGCCAG 632

Qy 621 ACCAGCAAGATCAAGATTTCAAGCTCTTGAAGATTCCTGCGCTAGAGATGTA 680
Db 633 ACCTGCTGAGTCGATTTCCAGGTGCTGAAATTTGCTCGAAAGTTGAAATGTAC 692

Qy 681 CCGGTTGCAACCGGCAAGAGCAAGGAGGAGCAAGATCAATCTGCGCTGCAAC 740
Db 693 CAGATTTCAATGCTTCTGACAGGAAAGAACCAAGATTCACATGGCAGTTTCCCAT 752

Qy 741 GGGAAATCTAGTGTTCAGGAGTTTCACTAAGATCAATGCTTCAACTGGGCCA 800
Db 753 GGGGTACTGTGTTCAGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 812

Qy 801 GAAAGCTGAGCTTCAAGAGGAAGCCCTTCTCATCAAGCTCGGCGCAAGTATG 860
Db 813 NNN 872

Qy 861 GTAACAGATACCTTGAATTCCTGATGCGCAGTGGGATTTCTGCAAGTCTTGGAA 920
Db 873 TTACAGAGACACATTAGAAATTTTGTGGTAGTAGATGAAATGTAAAGATCTTGGAA 932

Qy 921 AATCTGTGTGAACATCATGCTCTTTTGAAGTCTTTTGAAGAGCCCAACCAAGCCCAA 980
Db 933 GATTGTGTGAGATACACATTTTGTAGCTTTTGAAGCAACCTTAAGCAAAAGCAAA 992

Qy 981 GCCGCTCTTTTGAAGCGGGGTCATCTTGGTTCAAGTGTGCGACTCAGAGAGAGCT 1040
Db 993 AGCGCTTCTTCAAGCGGGGCTCTCTCAATACAGNNNNNNNNNNNNNNNNNNNN 1052

Qy 1041 TCTGACATATGTTAAAGAGAGACATAAGAGTGCAGTTTGAAGAGACAGCA 1100
Db 1053 NNN 1112

Qy 1101 GATTCAATTCATTCGAGACCTTCTTCAAGCTTCAAGAACTGAATTCGAAAGTCTGGA 1160
Db 1113 GACCCACAGCTC-----GTTGAGCTCTGACTGACAGACTTACCAAA 1154

Qy 1161 GCAGTCTCAGACAGACACAGCCTTACATTTGAGAGAGTGGCCGAATCTCCAGGGGCCA 1220
Db 1155 ACAGAGCATCTATTCCTGAGGAGATTGAGACTCTGCTCCCATCTTCAAGGAATGC 1214

Qy 1221 GAGCTGCGCGAGAGAAAGGAAACGGAAGTTTCGCGCGGAGACCGGGGTGCGACCCGAG 1280
Db 1215 CTTTACTGCTCTCTCTCCCTCACTCTGCTCCCTCTGCGCTGCGAGTTTAAAGACAG 1274

Qy 1281 CCTGCGCGAGAGAGAACCCCGGGTAAACAAGACGCGAGCGAGCTCTGCGGCC 1340
Db 1275 CAGAGCTCCTCAGACAGATCCCAAGTTTCTTCAATAGATCCAGCTCAGAGAGCG 1334

Qy 1341 CAGGAGAGAGAGAGAGAG---TCTTAAGATAGAGCCACAGAGTAACTCAGCC 1397
Db 1335 CAGTGAAGCAGTGTGAGAGGCCCAACACACATCGGCTCAGCCCTCGAGCCCGCC 1394

Qy 1398 CCGCAGCCCAAGACAGAGCTCCCTGACTGCGAGTCTCACTTTCGAGCTGTCTGAA 1457
Db 1395 ACTTCAGCTGTGTGAGGCTTTCAGAGAGATCTCTCAGCTTCTCCCTCAGCGGAA 1454

Qy 1458 CTGCGAGGGGAGTGGCCCTGCAACGTGACCTTGTCTCCAACTGAGCCCGAC 1517
Db 1455 GAGCCCTGAGTCTGAGCCCTGATTTCAAGTGCTT-----TGAGCCAGCTGA 1505

Qy 1518 CAAGCAGGCTCTCTCTTGAATCAAGCCGCTGCTGAATGACAGAGCTCCCGGACGGA 1577
Db 1506 ACAAGGCTCATCCCACTCTGAGCCCTGTCTCAGTGAAGTGTGGCGAGCCGGAGTGA 1565

Qy 1578 CGATGAGATGAGGCGCGAGAGAGATTCCCACTGATTAAGCGTACTTATAGCTAA 1637
Db 1566 CTGCGAG---GAGCCAGACACAAAGNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1622

Qy 1638 GGAAGTGTCTACACAGAGGAAATATCTGAAGATCTCGAAATTATCATCTTGTGTT 1697
Db 1623 NNN 1682

Qy 1698 TCAGAGCAGTGTAGAGAGAGAGAGCGCATGCGGAGAGCTGTAAGTCTCATATTC 1757
Db 1683 CCGCAGGCGAGTGTGAAGAGAGCGCATGCTGGAATCTTATGACGTGTCTTCTC 1742

Qy 1758 GAATTTGAACCTTTCACAAATTTTCACTAATTTTCTCAAGAAATTGAGCAAGACT 1817
Db 1743 CAACATCATCTCATATGATTTCAAGAGGCTTCTGCGGAGTGTAGAGAGAGCT 1802

Qy 1818 TGCCTGTGGAAGCGCGCTCAATATCAAG---ATTACCAAGAAATCGCGGA 1874
Db 1803 GGCACCTCTGGAGAGGCGCTCAAAAGCCACAAAGGCAATCAAGATTCGGGGA 1862

Qy 1875 TGTCTGTGGAAGAAATTCAGGAGCATGAAGCACTGCGGCTCACCTGTGAGAGACAG 1934
Db 1863 CATCTGTGAGAGACATGCGCATTAAGAGGATTTCAGAGTACTTCAAGAGACATGA 1922

Qy 1935 CGAGGCTTGTGAGGCGCTGAGAGATGAATCAAGAGCTCCGCGGCTGAGAGAACTTCTG 1994

```

Db      1923 CGAGGTCCTTAACGAACTGGAAAAAGGCTAACAAACGCTGTAAAGAAAGTTGAGGCAAGTGA 1982
Qy      1995 CAGAGACTTTGAGCTGAGAAAGGTGTGTTACTACCGCTCAACACCTTCTCTCGGAGCC 2054
Db      1983 CAAGGAAATTGAGCTGAGAAAGGTCTGTCTACTTGTCCCTCAACACGTTCTGTGTGAAGCC 2042
Qy      2055 ACTGCACCGGCTCATGCTACATCAAGCAGGTCTGTGAGCGGCTGTGCAACCAACCGGCC 2114
Db      2043 CATTCAGCGGCTGTGCTGACTACACCGCCCTGTCTGTCCGCCCTGTATCGGAGATTAACAGCCC 2102
Qy      2115 GAGCCACGCCGACTTCAAGGAGCTGCGAGCCGCTTTGGCAGAGATCACGAGATGTGTGC 2174
Db      2103 CGGGACATGACTAGCTAGCTGAGCCATGACGCCCTGAAGACATCAACAGAGTGAACAC 2162
Qy      2175 ACAGCTTCCAGGTAACATGATCAAGATGAGAAATTTCCAAACCTGACACAACTCAAGAA 2234
Db      2163 CACACTACAGCACATTTCTCATCCGGCTGAGAACTGTGCAAAAGCTAACGAGCTGCAGCG 2222
Qy      2235 AGATTGATTGGCAATTGACAAATCTTGGTTCGGGAAAGGAGTTTCATCCGTCTGGGAG 2294
Db      2223 GAGCTGTGTGGGCAATGAGAACTCATTTGCTCTGTGGAGGAGTTTCATCCGTGAGGCTG 2282
Qy      2295 CCTCAGCAAGCTCTCGGGAAAGGGGCTCCAGACAGCGATGTTCTCTGTTCACGACGT 2354
Db      2283 CTTTCAAGACTCACCAAGAAAGGGCTGCAGACAGAGATGTTTTTCTGTCTCAATAT 2342
Qy      2355 CCTGTCTATACAGACCGCGGGGCTGACGGCTCAATCATTTTAAAGTCCACGGGACGT 2414
Db      2343 GTTGTGTATACAAAGCAAAAGAGTGTGAGGAGCACGACCTTCGGATCGGGGCTCTCT 2402
Qy      2415 CCGGCTCTATGATGATGATGATGAGAGACGAAAGCAAGTGGGGGTGCCCCATGCT 2474
Db      2403 TCCCTCTCAAGGATGTGTGTGAGAAAGATGATTAAGATGTGTCTTTCACATGTTT 2462
Qy      2475 GACCTTCGGGGCCAGCGGACATCATGTGTGGCCGCAAGTTCTGTCTCCAGATGA 2534
Db      2463 CACCATCTACGCGGCTCAGAAACAAATCTGTGTGGCAGCAGACCTCGGCTGAGAAAG 2522
Qy      2535 GAAGTGGGTGAGAGACATCCAGATGGCCATGACCTGGCGGAGAAAGACGACGCCCGC 2594
Db      2523 GAAGTGAATGCTGAACTGAACTCCGGATTCACAGACGCAAGATGGCGGTGACACGGC 2582
Qy      2595 CCCTGAGTTCTGTGGCCAGCA-----GCCCCCTGCAACAAAGTCCCTGATGAACCAAC 2648
Db      2583 CCTGTGACCTGACGCGCGCATGTGTGCTCATCTGTCCGCCAATCCCCCAAGAGATATC 2642
Qy      2649 CGCGGCTGACCAAGAGTCAAGAGATGACTGACCGCTCGCGCAATCGCTGAGCGCCA 2708
Db      2643 TCTG---GAGCAGAGATCAAGATGATGTCTCGGGGTGTCCGCAAGCTCCCTGAGGCGCA 2699
Qy      2709 GGGCCCGGCAACCGGGGCAACAAATGATGACGTGTGTGCAACCGCAACCAAGCGCTTC 2768
Db      2700 TGGCCAGACACCGGAGCAACCAACCAATGACGTGTGTGTACCGGAACCAAGCGGTTC 2759
Qy      2769 CATGTGTGACTTGTAGATCGCAGTGTGAATTCAGTTGTCTGGAACCTGTGTGAGAAAT 2828
Db      2760 CAGGGCAGACCAAGTGTGAGCTGTCCAGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2819
Qy      2829 CAAAAACAGCAAGGTGTGCAAGAGCTGTGTGTGTTCACAACTTCTGTCTTCTT 2888
Db      2820 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2879
Qy      2889 CTACAAATCAACACAGGACAAATGATCCCTTTCAGCGCTGCTGTGTGGCTACTCGCT 2948
Db      2880 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2939
Qy      2949 CACCATCCCTCTGTAGTCCGAGAAATTCAGAAAGCTACGTGTTCAGAGTGTCACTTCAA 3008
Db      2940 GAGCATCCCAAGGAGGCGATGAGCAATACCAAAAGCTAGTTTCAAGTTCAGATTCAA 2999
Qy      3009 GTCCCACTTACTTCAAGGGCGAAGACGATACCTTGAAAGGTGATGAGAT 3068

```

```

Db      3000 ATCCACGCTACTTCTTCCGGGCTGAGAGCAAGTACATTTGAAAGGTGATGAGGT 3059
Qy      3069 GATCCGAGTGCACACAGCTGTGCGCGGAGACCC 3103
Db      3060 GATTCAGGGGGCCAGCACTCAGCGGGAGGCCC 3094

RESULT 14
BUS39080
LOCUS      959 bp      mRNA      linear      EST 13-SEP-2002
DEFINITION AGENCOURT_10215257 NIH_MGC_107 Homo sapiens cDNA clone
IMAGE:6569730 5', mRNA sequence.
ACCESSION BUS39080
VERSION    BUS39080.1 GI:22849521
KEYWORDS  EST.
SOURCE     Homo sapiens (human)
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.
REFERENCE  1 (bases 1 to 959)
AUTHORS   NIH-MGC http://mgi.nci.nih.gov/.
TITLES    National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
          Email: cgaabs-remail.nih.gov
          Tissue Procurement: ATCC
          cDNA Library Preparation: Rubin Laboratory
          cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
          DNA Sequencing by: Agencourt Bioscience Corporation
          Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LNL at:
          http://image.llnl.gov
          Plate: LLCM2757 row: h column: 18
          High quality sequence stop: 706.
FEATURES
source     1..959
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:6569730"
            /tissue_type="adenocarcinoma, cell line"
            /lab_host="DH10B (phage-resistant)"
            /clone_1ib="NIH MGC 107"
            /note="Organ: breast; Vector: pOTB7; Site 1: EcoRI;
            Site 2: XhoI; cDNA made by oligo-dT priming.
            Directionally cloned into EcoRI/XhoI sites using the
            following 5' adaptor: GGCACGAG(G). Library constructed by
            Ling Hong in the laboratory of Gerald M. Rubin (University
            of California, Berkeley) using ZAP-cDNA synthesis kit
            (Stratagene) and Superscript II RT (Life Technologies).
            Note: this is a NIH_MGC Library."
ORIGIN
Query Match      26.8%; Score 839.8; DB 3; Length 959;
Best Local Similarity 97.8%; Pred. No. 1.2e-210;
Matches 861; Conservative 0; Mismatches 18; Indels 1; Gaps 1;
Qy      811 TTCAAGAGAACGCTTTTCTCATCAAGCTCGGCGCAAGATGCCAATGTGCTTACCAAGAT 870
Db      1 TTCGGACAGAGCGCTTCTCTCAAGCTCCGGCCAGATGCCAATGTGCTTACCAAGAT 60
Qy      871 ACCTTGGAATTCCTGATGCGAGGATTTCTGCAAGTCTTCTGAAAAATCTGTGT 930
Db      61 ACCTTGGAATTCCTGATGCGAGGATTTCTGCAAGTCTTCTGAAAAATCTGTGT 120
Qy      931 GAACATCATGCTTCTTTTGAAGAGCCCAACCAAGGCCCAAGCCCGTCTC 990
Db      121 GAACATCATGCTTCTTTTGAAGAGCCCAACCAAGGCCCAAGCCCGTCTC 180
Qy      991 TTTAAGCCGGGGTCAATTTTGGTTCAAGTGTGAGTCAAGAGAGAGTTTCCACTAT 1050
Db      181 TTTAAGCCGGGGTCAATTTTGGTTCAAGTGTGAGTCAAGAGAGAGTTTCCACTAT 240

```

QY 1051 GTTAAAGAGAGACATTAAGAGTGCATTTGAAAGAGACAGCAAGATTCATCT 1110
DB 241 GTTAAAGAGAGACATTAAGAGTGCATTTGAAAGAGACAGCAAGATTCATCT 300
QY 1111 ATCCGAGCCTTGTCTTCAACCCCTACAGATTCGAAGTGTCTGAGCACTCTAG 1170
DB 301 ATCCGAGCCTTGTCTTCAACCCCTACAGATTCGAAGTGTCTGAGCACTCTAG 360
QY 1171 CAGAGCACACCTTACATTTTGAAGAGTCCGAATCTCCAGGAGGCGCAGAGTCTCCG 1230
DB 361 CAGAGCACACCTTACATTTTGAAGAGTCCGAATCTCCAGGAGGCGCAGAGTCTCCG 420
QY 1231 CGAGAGAAAGAAACGAAAGTTCCTCCGCGGAGGCGCGGAGTTCAGCCGAGCCGCG 1230
DB 421 CGAGAGAAAGAAACGAAAGTTCCTCCGCGGAGGCGCGGAGTTCAGCCGAGCCGCG 480
QY 1291 AGAGAAAGCCCGCGGTTAACAGACAGCGAGCCGCTCGGCGCCAGAGAGAA 1350
DB 481 AGAGAAAGCCCGCGGTTAACAGACAGCGAGCCGCTCGGCGCCAGAGAGAA 540
QY 1351 GAGAGAGAGTCTTAAAGATAGACCCAGAGATAAAGTCAAGCCCGCAGCCAGC 1410
DB 541 GAGAGAGAGTCTTAAAGATAGACCCAGAGATAAAGTCAAGCCCGCAGCCAGC 600
QY 1411 ACAGGCTCCCTGACCTGCGAGTCTCACTTCCGAGCTGTGTGAAGTCTGAGAGGGA 1470
DB 601 ACAGGCTCCCTGACCTGCGAGTCTCACTTCCGAGCTGTGTGAAGTCTGAGAGGGA 660
QY 1471 GTGAGCCCTGCGCAAGTGAAGTCTTCCCAAGTGAAGCCCGACCAAGAGGCTCT 1530
DB 661 GTGAGCCCTGCGCAAGTGAAGTCTTCCCAAGTGAAGCCCGACCAAGAGGCTCT 720
QY 1531 CCTTGTATCAGCCCGCTGCTGAATGACCAAGCTGCTCCCGAGAGAGATGAG 1590
DB 721 CCTTGTATCAGCCCGCTGCTGAATGACCAAGCTGCTCCCGAGAGATGAG 780
QY 1591 GGCGGAGAGAAAGATTCCTCAAGTGAAGCTTCAAGTGAAGTCTCAAG 1650
DB 781 GGCGGAGAGAAAGATTCCTCAAGTGAAGCTTCAAGTGAAGTCTCAAG 839
QY 1651 ACCGAGCAATATCTGAAGATCTCGAAGTATCACTT 1690
DB 840 ACCGAGCAATATCTGAAGATCTCGAAGTATCACTT 879

RESULT 15
BX350128/c 857 bp mRNA linear EST 08-APR-2004
LOCUS BX350128 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION clone CS0D1043YD12 3-PRIME, mRNA Sequence.
ACCESSION BX350128
VERSION BX350128.1 GI:30367428
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniota; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 857)
AUTHORS L.L.W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Pull-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
7229.f

For more information about this cluster, see
<http://www.genoscope.cns.fr/cdna?c=CS0D1043YD12>

FEATURES
source
Location/Qualifiers
1..857
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1043YD12"
/issue_type="PLACENTA COT 25-NORMALIZED"
/clone_id="Homo sapiens PLACENTA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
Query Match 26.6%; Score 834; DB 4; Length 857;
Best Local Similarity 99.8%; Pred. No. 3.9e-209;
Matches 856; Conservative 0; Mismatches 0; Indels 2; Gaps 2;
1689 TTGCTGTTTCAAGACAGTGAAGCAAGAGAGCCATCCCGAGACATGAAGTCT 1748
DB 856 TTGCTGTTTCAAGACAGTGAAGCAAGAGAGCCATCCCGAGACATGAAGTCT 797
QY 1749 CATATTCGCAATTTTGAACCTTTCGCAATTTTCAATTTTCTCAAGAAATTGA 1808
DB 796 CATATTCGCAATTTTGAACCTTTCGCAATTTTCAATTTTCTCAAGAAATTGA 737
QY 1809 GCAACGACTTGCCTGTGAGAGAGCGCTCAAAATGCCAATCAGAGATTACCAAAAGAT 1868
DB 736 GCAACGACTTGCCTGTGAGAGAGCGCTCAAAATGCCAATCAGAGATTACCAAAAGAT 678
QY 1869 CGGAGATGATCTGAAGAAATTCAGGAGATGAGACACCTGCGGCTCACTGTGAA 1928
DB 677 CGGAGATGATCTGAAGAAATTCAGGAGATGAGACACCTGCGGCTCACTGTGAA 618
QY 1929 GCAACGAGAGCCCTTGAAGAGCCCTTGAAGATGAATGAAGTCCCGGAGGCTGAGAA 1988
DB 617 GCAACGAGAGCCCTTGAAGAGCCCTTGAAGATGAATGAAGTCCCGGAGGCTGAGAA 558
QY 1989 CTTCGTGAGAGACTTGAAGTGAAGAGTGTGTACTTACCGCTCAACACTTCTCTCT 2048
DB 557 CTTCGTGAGAGACTTGAAGTGAAGAGTGTGTACTTACCGCTCAACACTTCTCTCT 498
QY 2049 GCGGCACTGCACTCGGCTCATGCACTTCAAGAGTCTTGAAGGCTGTGCAAAACCA 2108
DB 497 GCGGCACTGCACTCGGCTCATGCACTTCAAGAGTCTTGAAGGCTGTGCAAAACCA 438
QY 2109 CCGGCGAGAGCAAGCGCACTTCAAGAGATCCGAGCCGCTTGAAGATCCGAGAT 2168
DB 437 CCGGCGAGAGCAAGCGCACTTCAAGAGATCCGAGCCGCTTGAAGATCCGAGAT 378
QY 2169 GGTGAGCAGCTCCACGAGTGAAGTGAATCAAGATGAGAAATTTCCAGAGCTGACGAACT 2228
DB 377 GGTGAGCAGCTCCACGAGTGAAGTGAATCAAGATGAGAAATTTCCAGAGCTGACGAACT 318
QY 2229 CAAGAAAGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 2288
DB 317 CAAGAAAGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 258
QY 2289 GGGCAGCCTGAGAGAGTCTCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2348
DB 257 GGGCAGCCTGAGAGAGTCTCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 198
QY 2349 GCAAGTCTCTTATACAGAGCCGAGGAGTGAAGGAGTCCCAATCAGTTTAAATTCACAGG 2408
DB 197 GCAAGTCTCTTATACAGAGCCGAGGAGTGAAGGAGTCCCAATCAGTTTAAATTCACAGG 138
QY 2409 GCAAGTCTCTTATACAGAGCCGAGGAGTGAAGGAGTCCCAATCAGTTTAAATTCACAGG 2468
DB 137 GCAAGTCTCTTATACAGAGCCGAGGAGTGAAGGAGTCCCAATCAGTTTAAATTCACAGG 78

QY 2469 CTGCTGACCCCTCGGGGCCAGGCGACATCATGTGCGCCGAGTTCTCGTCCGA 2528
Db 77 CTGCTGACCCCTCGGGGCCAGGCGACATCATGTGCGCCGAGTTCTCGTCCGA 18
QY 2529 GATGAGAAAGTGGGTTGA 2546
Db 17 GA-GAAGAAAGTGGGTTGA 1

Search completed: May 22, 2006, 08:59:33
Job time : 13995.1 secs

This Page Blank (uspto)

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 22, 2006, 00:49:20 ; Search time 542.44 Seconds
(without alignments)
10813.957 Million cell updates/sec

Title: US-09-555-342b-1_COPY_49_3183

Perfect score: 3135
Sequence: 1 atggggaatagatagacagag.....acaagagctctctgtatc 3135

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1403666 seqs, 935554401 residues

Total number of hits satisfying chosen parameters: 2807332

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Issued Patents NA:*

- 1: /EMC_Celerra_SIDS3/prodata/2/ina/1_COMB.seq:*
- 2: /EMC_Celerra_SIDS3/prodata/2/ina/5_COMB.seq:*
- 3: /EMC_Celerra_SIDS3/prodata/2/ina/6A_COMB.seq:*
- 4: /EMC_Celerra_SIDS3/prodata/2/ina/6B_COMB.seq:*
- 5: /EMC_Celerra_SIDS3/prodata/2/ina/7_COMB.seq:*
- 6: /EMC_Celerra_SIDS3/prodata/2/ina/H_COMB.seq:*
- 7: /EMC_Celerra_SIDS3/prodata/2/ina/PC/US_COMB.seq:*
- 8: /EMC_Celerra_SIDS3/prodata/2/ina/RE_COMB.seq:*
- 9: /EMC_Celerra_SIDS3/prodata/2/ina/RE_COMB.seq:*
- 10: /EMC_Celerra_SIDS3/prodata/2/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	204.2	6.5	310	US-09-621-976-8193	Sequence 8193, Ap
2	192.4	6.1	2853	US-09-949-016-1933	Sequence 1933, Ap
3	180.4	5.8	6263	US-09-664-958-5	Sequence 5, Appl1
4	180.4	5.8	6268	US-09-566-921-57	Sequence 57, Appl1
5	169.6	5.4	4292	US-09-949-016-1488	Sequence 1488, Ap
6	169.6	5.4	4236	US-09-949-016-208	Sequence 208, Ap
7	165.8	5.3	576	US-09-270-767-1995	Sequence 1995, Ap
8	165.8	5.3	576	US-09-270-767-1995	Sequence 1995, Ap
9	144	4.6	3984	US-09-848-294-1	Sequence 1, Appl1
10	143.2	4.6	2872	US-09-906-779-3	Sequence 3, Appl1
11	142.4	4.5	2156	US-09-949-016-3593	Sequence 3593, Ap
12	133.6	4.3	3398	US-09-799-451-201	Sequence 201, Ap
13	106.4	3.4	449	US-09-270-767-1542	Sequence 1542, Ap
14	106.4	3.4	449	US-09-270-767-1542	Sequence 1542, Ap
15	105.2	3.4	184	US-09-621-976-15728	Sequence 15728, A
16	97.8	3.1	17098	US-09-949-016-15864	Sequence 15864, A
17	95	3.0	2493	US-10-104-047-41	Sequence 41, Appl1
18	90.6	2.9	573	US-09-270-767-165	Sequence 165, Appl
19	90.6	2.9	573	US-09-270-767-165	Sequence 1544, A
20	65.2	2.1	3438	US-10-164-595-29	Sequence 29, Appl1
21	64	2.0	4272	US-09-949-016-5747	Sequence 5747, Ap
22	63.8	2.0	1626	US-09-620-312D-1033	Sequence 1033, Ap
23	63.8	2.0	1699	US-09-913-171A-26	Sequence 26, Appl1

24	62.8	2.0	2730	US-09-774-528-175	Sequence 175, App
25	62.8	2.0	2730	US-10-120-988-175	Sequence 175, App
26	62.8	2.0	3503	US-10-094-749-560	Sequence 560, App
27	59.6	1.9	722	US-09-270-767-1861	Sequence 1861, App
28	59.6	1.9	722	US-09-270-767-1861	Sequence 1861, App
29	59.4	1.9	222	US-09-016-434-45	Sequence 45, Appl1
30	56.8	1.8	4080	US-08-446-345-35	Sequence 35, Appl1
31	56.8	1.8	4081	US-09-949-016-1292	Sequence 1292, Ap
32	53.8	1.7	7218	US-08-232-463-14	Sequence 14, Appl1
33	50.2	1.6	3044	US-09-814-915A-84	Sequence 84, Appl1
34	50.2	1.6	3044	US-09-880-107-3718	Sequence 3718, Ap
35	48.8	1.6	1308	US-09-252-991A-11936	Sequence 11936, A
36	48.8	1.6	1335	US-09-252-991A-12073	Sequence 12073, A
37	48.2	1.5	1816	US-09-620-312D-1439	Sequence 439, App
38	48.2	1.5	3415	US-09-902-540-5107	Sequence 5107, App
39	48.2	1.5	28058	US-09-902-540-5107	Sequence 1252, Ap
40	47.4	1.5	2400	US-08-930-001-1	Sequence 1, Appl1
41	47.4	1.5	2400	US-09-091-885-1	Sequence 1, Appl1
42	47	1.5	801	US-09-902-540-8879	Sequence 8879, Ap
43	47	1.5	12028	US-09-902-540-1071	Sequence 1071, Ap
44	47	1.5	54711	US-09-949-016-17489	Sequence 17489, A
45	46	1.5	235452	US-09-949-016-13675	Sequence 13675, A

ALIGNMENTS

```
RESULT 1
US-09-621-976-8193
; Sequence 8193, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 8193
; LENGTH: 310
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-8193

Query Match          6.5%; Score 204.2; DB 3; Length 310;
Best Local Similarity 98.1%; Pred. No. 1.8e-44;
Matches 203; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY      1  ATGGGAAATAGAGCAGAGCCGACCCAGATCGACTGGGGCCCCGGAAATTGCG 60
      |||
DB      104 ATGGGAAATAGAGCAGAGCCGACCCAGATCGACTGGGGCCCCGGAAATTGCG 163

QY      61  GGGATCAGTACCTTGGAGTGAAGCGCCGCCCAACCTTCAGAAATCGTG 120
      |||
DB      164 GGGATCAGTACCTTGGAGTGAAGCGCCGCCCAACCTTCAGAAATCGTG 223

QY      121 TCCATCAATATCAGATGTCGATGACACCCAGAGCAATTTGAAATTCAGAAAGCT 180
      |||
DB      224 TCCATCAATATCAGTATGTCGATGACACCCAGAGCAATTTGAAATTCAGAAAGCT 283

QY      181 CCTGGAGAGTGTCTGTGATGCAATT 207
      |||
DB      284 CCTGGAGAGTGTCTGTGATGCAATT 310

RESULT 2
US-09-949-016-1933
; Sequence 1933, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
```



```

?  APPLICANT: VENTER, J. Craig et al.
?  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
?  TITLE OR INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
?  FILE REFERENCE: CLO01307
?  CURRENT APPLICATION NUMBER: US/09/949,016
?  CURRENT FILING DATE: 2000-04-14
?  PRIOR APPLICATION NUMBER: 60/241,755
?  PRIOR FILING DATE: 2000-10-20
?  PRIOR APPLICATION NUMBER: 60/237,768
?  PRIOR FILING DATE: 2000-10-03
?  PRIOR APPLICATION NUMBER: 60/231,498
?  PRIOR FILING DATE: 2000-09-08
?  NUMBER OF SEQ ID NOS: 207012
?  SOFTWARE: FastSeq for Windows Version 4.0
?  SEQ ID NO 1933
?  LENGTH: 2853
?  TYPE: DNA
?  ORGANISM: Human
?  US-09-949-016-1933

```

Query Match	6.1%;	Score 192.4;	DB 3;	Length 2853;
Best Local Similarity	52.8%;	Pred. No. 7.8e-41;		
Matches 490;	Conservative 0;	Mismatches 426;	Indels 12;	Gaps 3;

Qy	117	CGTGTCCATCAAAATCCAGATGCTGGATATGACACCCAGAGGCACTTGAAGTTCACAAG	176
Db	784	CATGACCTGCAGAGTTCTTTGTTGGATGACACAGTTTATGATGTGTGGAGAAACA	843
Qy	177	AGTCTCGGGAGGTGCTGCTGATGACAGTTTGCAACCACTCAACTCGTGAAGTGA	236
Db	844	TGCTAAGGGACAAAGATTGTGCTTAAACGATATGTAGAGATCTCAATCTTTTGGAAAGAGA	903
Qy	237	CTATTTTGGGCTCGAGTTTCCGTATCACAAAAAGATCAGCGGTGGGCTGATCTCTAAA	296
Db	904	CTATTTTGGTCTTACCACTTTGGGATTAACGCACTCTTAGACATGGCTGATTTCCGCCAA	963
Qy	297	ACCATTGTGTAAACAGATTAGAAAGCCAAAGCAGTTGTTGTTAAGTTGTGTAAATT	356
Db	964	AGAAATTAAGAAAGCAGGTCGTGG---TGTCCTTGGAATTTTACATTAATGTAAGATT	1020
Qy	357	CTTTCCGCGCTGACACACACAACTCCAAAGAACTCACAAAGTACCTGTTCCGGCTGCA	416
Db	1021	TTATTCACCTTGACCAGACAGTTTAACAGAAACATTAACAAGATTAATTAATCTCTTCA	1080
Qy	417	GGTGAACAGGACTTGTGCTCAAGCAGGTTGACGTTAATGACACACAGCGACGCTCTT	476
Db	1081	GCCTTCGACAGGACATAGTTGACGAGACGCTGCTGTTCCTTTGCAACCTTAGCATTAAT	1140
Qy	477	GATTTCACACTTGTGCATCTTGAGATTGGGGATTTTGTATGTAACCTTGGACACAGAGCA	536
Db	1141	AGGTTCTTTCACCACTTCACAGTCTGAACTGGGAGACTAAGACCCAAACCTCCATGGCGTGA	1200
Qy	537	CTTTCGAAAAAATAATATACATACC-----TCAGCAAGAAGCACTAGAGCAAAATCGT	590
Db	1201	TTATGTTTAGTGATTTTAACTGCGCCCGGAATCABACCAAGAACTTGAAGAGAGTCAT	1260
Qy	591	GGAAATTCAACCATTAACAACATTGSAACAAACACAGCAGATCAGATTTCCAGCTCTAGA	650
Db	1261	GGAACTCATTAAGTCATACAGGTCATCAGCTACGCTCAGGCTCAGCTTAGGAGTTTCCTGA	1320
Qy	651	GATGACCGCTGCGCTAGAGATGTATGGAATCCGGTGAACCCGGGCAAGACAGGAGAG	710
Db	1321	GAATGCCAAAAAGTTGTCTATGTATATGAGATTGATCTTCAATAAGCAAAAGACTTGGAAAG	1380
Qy	711	CACGAGATCAATCTGGCCGTTGCAACACGGGAATTCATGTGTTTCAAGGTTTCACTAA	770
Db	1381	AGTAGATATCATCTTAGGTGTCTCTCTAGTGGCCCTTCTGGTTTAAAGATTAAGCTGAG	1440
Qy	771	GATCAATGCGCTTCAACTGGGCGCAAGGTGGGAGACTGAGCTTCAAGAGGAAGCGCTTCT	830
Db	1441	AATTAACCGCTCTCCCTTGGCCCAAGGTGTAAGATTTCTTATATAAGTATAGACTTTTT	1500
Qy	831	CATCAAGCTCCGCGCAGATGCAATAGTGCCTAACAGATACCTTGGAAATTCCTGATGGC	890

Db	1501	CATTAAATTCGGCGCTGAGAGCAAGACAGTATGAATGACATCGGATTCAAACTCC	1560
Qy	891	CAGTCGGGATTTCTGCAGTCCTTCTGGAAATCTGTGTGAACATCATGCTTCTTAG	950
Db	1561	CAGTTACCGAGCAGCTAAGAAATTTATGGAAGCTGTGTGAACATCACAGTTTTTCAG	1620
Qy	951	ACTTTTGAAGAGCCCAACCAAGCCCAAGCCCGTCTCTTTTAGCCGAGGATCATATT	1010
Db	1621	A---TTGACATCTACAGACACACATTTCCCAAAAGCAAAATTTCTTGCGGTAGATCAAAATT	1677
Qy	1011	TCGGTTCAGTGTGTGGACTCAGAAGCAG	1038
Db	1678	TCGATACAGTGGCCGAGCTCAAGCTCAG	1705

RESULT 3
US-09-664-958-5
; Sequence 5, Application US/09664958

GENERAL INFORMATION:

APPLICANT: Canfield, Robert

APPLICANT: Rudchenko, Serge

FILE REFERENCE: 0575/60240
CUMMANT ADDITION NUMBER: 116/06/664 058

CURRENT FILING DATE: 2000-09-18
NUMBER OF SEC. 1D NOS.: 38

```

; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5

```

LENGTH: 6263
TYPE: DNA

ORGANISM: Human
; IIS-09-664-958-5

US-09-664-958-5

Query Match	5.8%	Score 180.4	DB 3	Length 6263
Best Local Similarity	52.9%	Pred. No. 1.9e-37		
Matches 485	Conservative	0	Mismatches 416	Indels 15
				Gaps 4

QY	13	GGCTGAGATGACACCCACAGAGGCACTTTGAAAGTTCCACAAGAGGCTCTGGGAAAGGCTGCG	196
Db	469	TGCTTGATGCTCCGGAGTAGTAGTGTAGGTTGAGGAAACATGCGCGGGGCCAGGTGCTGT	528
QY	197	TGATGACAGTTTGGCAACCACTCAACCTCGTGAAGGTGACTATTTTGGCCTCGAGTTTC	256
Db	529	TTGACCTGTGCTGTGAAACCTCAACCTCTTAGAAGAAAGACTACTTGGCGGCTGACCTTCT	588
QY	257	CTGATCACAAAAGATCAAGGTGTGGCTGTGATCTTCTTAAAAACCATTTGTGAAACAGATTA	316
Db	589	GTTGATGTGACAGCCAGACAGAAAGAACTGGGTGTGACCCCTCCAGAGATCAAGAGACGATCC	648
QY	317	GAAGGCCAAAGCAGTGTGTTGTTAAGTTTGTGGTGAATTTCTTCGCGCTGACCAACAC	376
Db	649	GGAG---TAGCGCCCTGGAAATTTTGCTTTCACAGTCAAGTTCTTACCGGCTGATCTGTGCC	705
QY	377	AATCTCAAGAAAGATCAACAAGTACTGTTGCGCGCTGCAGGTGAAGCAGACCTTGGCTC	436
Db	706	AGCTGACAGAAAGATCAACAAGATACTAACCTGTGCTGCAGCTCGCGGCACACATCATCA	765
QY	437	AAGCAGAGTTGACGTGTAAATGACACCAAGCGAGCTCTTGATTTCAACATTTGGCAAT	496
Db	766	CGGGCCGGCTGCAATGCTCTTTGTACAGCATGCGCTTCACTGGGCTCTTACGCTGTGACGG	825
QY	497	CTGAGATTTGGGATTTTGTATGAAAGCCTTGGACAGAGACCTTAGACAAAATAAATAACA	556
Db	826	CTGAGCTGGGGTACTATGATGTGTGAGGAGCATGTGGGCACTTATGTGACGACGCTCGGCT	885
QY	557	TACC-----TCAGCAAGACGCACTAAGAGCAAAATCTGTGAATTTCAACATAACACA	610
Db	886	TCGCGCCCTTAACCAACCCGGGAGCTGAGAGGAGAGATCATGAGACTGTGATTAAGCATATA	945

Qy	6111	TTGAGCAAAACACACACAGAAATCAGATTTTCAGCTCCTGAGATTTGCGCGTGCGTAAGGA	670
Db	946	GGGGGATGACCCCCGGGAGAGCAGAAATCTCACTTTTAGAGATGCCAAGAGCTTTTCCA	10055
Qy	671	TGTATGAAATCCGGTTGCAACCCGGCCAAAGACAGGAAAGGACGAAATCAATCTGCGCG	730
Db	1006	TGTACGGAGTGAACCTGCACCACTGCACCAAGAACTTGAGGGCATGCATCATGTTAAGCG	1065
Qy	731	TTGGCAACACGGGAATTCCTAGTGTTCACGGGTTTCACATMAATCAATGCCCTTCAATGG	790
Db	1066	TTTGTCCTAATGGCCTGCTCATCTAACCGGAGCCGCTGAGATCAACCGCTTTGCTGGC	1125
Qy	791	CCAAAGTCGGAGAGCTGAGCTTCAAGAGAAAGCGCTTCTCATCAAGCTCCGGCCACAGATG	850
Db	1126	CCAAAGTCTCTCAAGATCTCTCTCAAGAGAGAGATTAATCTTATATCAAGATCCGGCCTGGG	1185
Qy	851	CCAAATAGTGCCTACCAAGATACCTTTGAATTCCTGATGCGCAATCGGAGATTTCTGCAAGT	910
Db	1186	AGTATGAGCAATTTGAGAGCACAATTTGGCTTTAAGCTCCCAAAACCAACCGGTACGCAAGA	1245
Qy	911	CCTTCTGGAATCTGTGTGTGACATCAATGCGCTTTTAGACTTTTGAAGAGCCCAAG	970
Db	1246	GACTGTGGAAGTCTGCATTCAGCATATACATTTTCCGCG--TGGTGTCCCTTGAGC	1302
Qy	971	CAAAGCCCAAGCCCGCTCTCTTTTAGCCGGGGGTATCAATTCGGTTCAAGTGGTGGACTC	1030
Db	1303	CCCCACCCCAAGGGCTTCTGTGGA---TGGGCTCAAGTTCCGGTACAGTGGAGAGACC	1359
Qy	1031	AGAGACAGGTTCTCGA	1046
Db	1360	AGGCACAGACTCGCA	1375

```

RESULT 4
US-09-566-921-57
Sequence 57, Application US/09566921
Patent No. 668288
GENERAL INFORMATION:
APPLICANT: Loring, Jeanne F.
APPLICANT: Tingley, Debora W.
APPLICANT: Edwards, Carla M.
TITLE OF INVENTION: GENES EXPRESSED IN ALZHEIMER'S DISEASE
FILE REFERENCE: PA-0024 US
CURRENT APPLICATION NUMBER: US/09/566,921
CURRENT FILING DATE: 2000-05-05
NUMBER OF SEQ ID NOS: 138
SOFTWARE: PERL Program
SEQ ID NO 57
LENGTH: 6268
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: m1sc_feature
OTHER INFORMATION: incyle ID No. 6682888 453592.3
NAME/KEY: unsure
LOCATION: 2956, 4230-4252, 4825-4887
OTHER INFORMATION: a, t, c, g, or other
US-09-566-921-57

```

Query Match	5.8%;	Score 180.4;	DB 3;	Length 6268;
Best Local	Similarly	Pred. No. 1.9e-37;		
Matches 485;	Conservative	0;	Mismatches 416;	Indels 15;
				Gaps 4;

QY TGCTGAATGACACCCAGGAGCATTTGAAGTTCCACAAAGAGCTCTGGGAAGAGTGTCTG 196
 Db TGCCTTGATGCTCTGGAGTATGAGTGTAGAGTGTGAGAAACATGCGCGGAGCCAGTGTCTGT 528
 QY TGGATGACGTTTGCACACCACTCAACCTCGTGGAAAGTGACATATTTTGAGCTCGAGATTTC 256
 Db TTGACCTGCTGTGTACACCTCAACCTCTCTGAGAGAGATCACTTCCGAGCTGACCTTCT 588
 QY CTGATCACAAAAGTCAAGGTGTGGCTGTGATCTCTTAAACCACTTGTGAACGATTA 316

Db 589 GTGATGCTGCAGACGCAAGAAACTGGCTGAGACCCCTCCAAAGAGATCAAGAACGATCC 648

Qy 317 GAGGCCAAGCAGCTGTTGTTAAAGTTGTGTGAATTTCTTCCGCTGCACACAC 376

Db 649 GGAG---TAGCCCCCTGGAATTTTGGCTTACAGATCAAGTTCTTACCCGCTGATCTGCCCC 705

Qy 377 AACTCCAAAGAAATCCACAAGATACCTGTTCGCGCTGAGGGGAAGCAGGATTTGSGCTC 436

Db 706 AGCTACAGAAAGACATCACAGATATCTACTGTGCTGAGCTGCGGGCAGACATCATCA 755

Qy 437 AAGCAGGTTGACGTGTAATGACACCAAGCAGCTCTCTGATTTTACACATTTGSCAAT 496

Db 766 CGGCGCGGCTGCATGCTCTTTGTTCAGCAGTCCCTTACTGGGCTCTACGCTGTGAGG 825

Qy 497 CTGAGATTTGGGGATTTTGTATGAGCCTTGCACAGAGACACTTACGAAAAATTAATACA 556

Db 826 CTGACTGGGTACATATGATGCTGAGGAGACATGTGGGCAACTATGTACAGGACCTCGCT 885

Qy 557 TACC-----TCAGCAAGACGCACTAGAGAGCAAAATCGTGGAAATTTCAACATACACA 610

Db 886 TCGCCCTTACCAAGACCCGGGAGCTGGAGAGAGATCATGGAGCTGCATTAAGACATATA 945

Qy 611 TTGACAAACACACAGCAGATCAGATTTTCAGCTCTTAGAGATTGCCCCGCTAGAGA 670

Db 946 GGGGGATGACCCCGGAGAGCAGAAATTCACATTCTTAGAGAAATGCCAAGAACTTTCCA 1005

Qy 671 TGTATGGAATCCGGTTTGACACCCGGCCAAAGACAGGGAAGGCACGAATCAATCTGCCC 730

Db 1006 TGTATGGAGTGAACCTTGACCAATGCCAAAGACTCTGAGGGCAATCGCATCATGTTAGGCG 1065

Qy 731 TTGCCAACACGGGAATTTAGTGTTCAGGGTTTCACTAAGATCAATGCTTCAACTGGG 790

Db 1066 TTGTGGCCAAATGSCCTGCTCATCTACCGGACCCGCTGAGATCAACCCGTTTGCTGGC 1125

Qy 791 CCAAGTSCGGAAGCTGAGCTTCAAGAGGAAGCGCTTCTCATCAAGCTCCGGCCAGATG 850

Db 1126 CCAAGATCTCAAGATCTCTTCAAGAGAGGATTAACCTTCTATATCAAGATCCGGCTGGGG 1185

Qy 851 CCAATATAGTCGTAACAGATACCTTGGAAATCTCTGATGGCCAGTGGGATTTCTGGAAT 910

Db 1186 AGTATGAGCAATTTGAGAGACAAATTTGGCTTTAACTCCCAACCAACCGGTACGCCAAGA 1245

Qy 911 CCTTCTGGAATAATCTGTGTGAACATCAATGCTCTTTTGAAGCTTTTGAAGAGCCAAAC 970

Db 1246 GACTGTGGAAGTCTGCATCGAGCATCAATCTTCCGGC---TGTGTCCCTCGAGC 1302

Qy 971 CAAAGCCCAAGCCCGTCTCTTTAGCCGGGGTCAATCTTCGTTCAAGTGTGCGACTC 1030

Db 1303 CCCCACCCAAAGGCTTCTGTGTA---TGGGCTCAAGTTCCGGTACAGTGGAGGAGACC 1359

Qy 1031 AGAAGCAGTCTTCTGA 1046

Db 1360 AGGCACAGACTTCGCCA 1375

RESULT 5
 US-09-949-016-1488
 ; Sequence 1488, Application US/09949016
 ; Patent No. 6812339
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 ; WITH HUMAN DISEASE. METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CLO01307
 ; CURRENT APPLICATION NUMBER: US/09/949,016
 ; PRIOR FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: 60/241,755
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/237,768
 ; PRIOR FILING DATE: 2000-10-03
 ; PRIOR APPLICATION NUMBER: 60/231,498
 ; PRIOR FILING DATE: 2000-09-08
 ; NUMBER OF SEQ ID NOS: 207012

SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO: 1488
LENGTH: 4292
TYPE: DNA
ORGANISM: Human
US-09-949-016-1488

Query Match 5.4%; Score 169.6; DB 3; Length 4292;
Best Local Similarity 51.4%; Pred. No. 1.2e-34;
Matches 472; Conservative 0; Mismatches 434; Indels 12; Gaps 3;

```
QY 127 AAAATCCAGATGCTGATGACACCCAGAGGCACTTTGAAGTTCACAAAGAGCTCTGGG 186
DB 675 AAAGTACCCCTCTTAATGATGACCGAATCAGCTGACCTGGAGAAATGATGCAAGGGA 734
QY 187 AAGGTCTGCTGGATGACAGTTTGCACCCCTCAACCTGTGGAGAGTACTATTTTGGC 246
DB 735 CAAGTGTATTGTAACAAGTGTGAAACCTCAATCTCTTGGAGAAAGACTACTTTTGA 794
QY 247 CTCGAGTTTCTGATCACAATAAAGATCAGGTGTGGCTGGATCTCTAAACCATTTG 306
DB 795 CTTTGTCTCAGAAAGCCCTGACGAAATACTGTTAGTCTCTGTAAGAAATAAG 854
QY 307 AAACAGATTAGAAAGCCAAAGCACTTGTGTTAAGTTGTGTAATTTCTTCGCT 366
DB 855 AGCACTAGAAAC---TTCATGGCTATTCACTTTAATGTGAAGTTTATCTCT 911
QY 367 GACCAACAACAATCCAAAGAACTCACAAGTACTCTGCGGTGAGAGTGAAGCAG 426
DB 912 GATCTCTTCAATGACTAAGATATACCAAGTACTCTGTGCTTCAAGCTCGGCGAG 971
QY 427 GACTTGGCTCAAGGCAAGTGAAGTGAATGACACCAAGGCAAGCTCTTGTGATTTGACAC 486
DB 972 GACATTTGCTCTGCGCGCTGCTGCTCTTTTGTGATCATGATCTCTCTGGATCTTAC 1031
QY 487 ATTGTGCAATCTGAGATTGGGGATTTT-----GATGAAGCTTGTGACAGAGACCTTA 540
DB 1032 ACCCTGACAGCTGAACTTGGTGACTATGACCCAGAAAGAACTGGACATCGACCTCAGT 1091
QY 541 GCAAAAAATTAATATACCTACCTGAGAAAGCGACTGAGAGACAAATGATGGAATTTGAC 600
DB 1092 GAATTCAGTTGGCCCTTACTCAGACTAGAGAGCTGAAAGAGAGGTGAGAGCTGCAC 1151
QY 601 CATTAACCAATTTGACAAACACCAAGCAAGATCAGATTTCCAGCTCTTAGAGATTGCCGT 660
DB 1152 AAAACCAAGAGGGCTTATGCGACAGCAAGCTGATTTCCAGTTCTTAAAGAAATGCAAG 1211
QY 661 CGGCTAGAGATGTATGGAATCCGGTTGACCCGGCCAAAGGACAGGAGACAGAAAGATC 720
DB 1212 AGGCTTTCCATGATGATGTGTTGACCTAATGATGCAAGAGACTCAGAAAGGTGGACATC 1271
QY 721 AATTCGCGCGTTGCCAAACGCGGAATTTCTAGTTCAGGGTTTCACTAAGATCAATGCGC 780
DB 1272 AAGCTGGGGCTGTGTGCTAATGAGCTTCTCATTTACAAAGACAGACTGCGAATCAATCGT 1331
QY 781 TTCAATCGGGCCCAAGGTGCGGAGCTGAGCTTCAAGAGAAAGCGCTTCTCATCAAGCTC 840
DB 1332 TTTCGTTGGCGGAAATCTTTAAATTTCTTAATAAGCAGTAACTTCTTAATTAAGTC 1391
QY 841 CGGCGAGATGCAATATGTCGTAACAGATACCTTGAATCTCTGATGCGCAGTGGGAT 900
DB 1392 AGACCGCGAGAGCTGGAACAGTTTGAAGATACCATTTGAACTGCGCAACACCGG 1451
QY 901 TTCTGCAATGCTTTCGAAATCTGTGTTGAACATCATGCTCTTTTGAAGCTTTTGA 960
DB 1452 GCAAGCAAAAGACTATGAAAGTGTGCGTGAAGCATATATCTTTCAAGGCTGTGTT--- 1508
QY 961 GAGCCAAACCAAGGCCCAAGCCGTCCTTTAGCCGCGGCTCATCTTGGTTCAGT 1020
DB 1509 TGTCCAGAGAGCAACAAAGCCAAAGTTCCTGACCTTGGGCTCAAAATTTGGCTATAGT 1568
QY 1021 GGTGGACTCAGAGCAG 1038
```

DB 1569 GGCCGACCCAAAGCAG 1586

RESULT 6
US-09-949-016-208
Sequence 208, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
PRIOR FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR FILING DATE: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO: 208
LENGTH: 4336
TYPE: DNA
ORGANISM: Human
US-09-949-016-208

Query Match 5.4%; Score 169.6; DB 3; Length 4336;
Best Local Similarity 51.4%; Pred. No. 1.2e-34;
Matches 472; Conservative 0; Mismatches 434; Indels 12; Gaps 3;

```
QY 127 AAAATCCAGATGCTGATGACACCCAGAGGCACTTTGAAGTTCACAAAGAGCTCTGGG 186
DB 705 AAAGTACCCCTCTTAATGATGACCGAATCAGCTGACCTGGAGAAATGATGCAAGGGA 764
QY 187 AAGGTCTGCTGGATGACAGTTTGCACACCTCAACCTCTGTGAAGTGAATTTTGGC 246
DB 765 CAAGTGTATTGTAACAAGTGTGAAACCTCAATCTCTTGAAGAAAGACTACTTTTGA 824
QY 247 CTCGAGTTTCTGATCACAATAAAGATCAGGTGTGGCTGATCTCTAAACCATTTG 306
DB 825 CTTTGTCTCAGAAAGCCCTGAGCAAAAACTGGTTAGATCTGCTAAGAAATAAG 884
QY 307 AAACAGATTAGAAAGCCAAAGCAAGCTGTTGTTAAGTTGTGAGAAATTTCTTCGCT 366
DB 885 AGCACTAGAAAC---TTCATGGCTATTCACTTTTAATGTAAGTTTATCTCTCT 941
QY 367 GACCAACAACAATCCAAAGAAACTCAAGAGTACTGTTGCGCTGCAAGTGAAGCAG 426
DB 942 GATCTCTCAATGATGAGATATACCAAGATCTCTTGTGCTTCAAGCTCGGCGAG 1001
QY 427 GACTTGGCTCAAGGAGTGAAGTGAATGACACCAAGCAAGCTCTCTTGAATTTGACAC 486
DB 1002 GACATTTGCTCTGCGCGCTGCTGCTCTTTTGTGATCATGCTCTCTGGAGATCTAC 1061
QY 487 ATTGTGCAATCTGAGATTGGGGATTTT-----GATGAAGCTTGTGAGAGAGACCTTA 540
DB 1062 ACCCTGACAGCTGAATCTTGTGATGATGACCAAGAAACATGAGCATGACCTCAGT 1121
QY 541 GCAAAAAATTAATATACCTCAGCAAGACGCACTGAGAGCAAAATGTTGAATTTTAC 600
DB 1122 GAATTCAGTTTGGCCCTTACTCAGACTAAGAGCTGGAAGAGAGTGGCAGAGCTGCAC 1181
QY 601 CATTAACCAATTTGACAAACACCAAGAAATCAAGATTTCCAGCTCTTGAAGTTGCCGT 660
DB 1182 AAAACCAAGAGGGCTTATGCGCAAGCAAGCTGATTTCCAGTTCTTGAAGAAATCAAG 1241
QY 661 CGGCTAGAGATGTATGGAATCCGGTTGACCCGCGCAAGGACAGGAGAGCAGAAAGATC 720
DB 1242 AGCTTTCCATGATATGATGTTGATGACTTATGATGCAAGAGCTCAGAAAGGTGTGACATC 1301
QY 721 AATTCGCGCGTTGCCAAACGCGGAATTTCTAGTGTTCAGGGTTTCACTAAGATCAATGCC 780
```

```
Db 1302 AAGTGGGCGGTGTGTCTTAATGAGCTTCTCATTTTCAAAAGACAGACTGCGAATCAATCGT 1361
Qy 781 TTCAACTGGGCCAAGGTGCGGAGCTGAGCTTCAAGAGGAGCGCTTTCATCAACCTC 840
Db 1362 TTGCTTGGCCGAAAATCTTAAAAATTTCTATTAAGCAGTAATCTTCAATTAAGTGC 1421
Qy 841 CGCCGAGATGCCAATAGTGGGTACAGAGATACCTTGAATTTCTGATGGCAGTGGGAT 900
Db 1422 AGACCGGCAAGCTGGAACAGTTGAGATACCAATGATTTCAACGCGCAACACCGG 1481
Qy 901 TTCTGCAAGTCTTCTGAAAATCTGTGTAACATCATGCTTCTTTAGACTTTTGA 960
Db 1482 GCAGCGAAAAGACTATGAAAAGTGTGCTGAGCATATCTTTTCAAGGCTTTGTT--- 1538
Qy 961 GAGCCCAACCAAGCCCAAGCCGCTCTTTAGCCGGGGTCAATTTCCGTTCACT 1020
Db 1539 TCTCCAGAGCAGCAACCAAGCAAGTCTTGAATTTGGGGTCCAAATTTCCGTTAGT 1598
Qy 1021 GGTCCGACTCAGAACG 1038
Db 1599 GCGCCGACCCAGACAG 1616
```

RESULT 7
US-09-270-767-1995
Sequence 1995, Application US/09270767

Patent No. 6703491
GENERAL INFORMATION:
APPLICANT: Homburger et al.
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
FILE REFERENCE: File Reference: 7326-094
CURRENT APPLICATION NUMBER: US/09/270.767
NUMBER OF SEQ ID NOS: 62517
SOFTWARE: Patent Ver. 2.0
SEQ ID NO 1995
LENGTH: 576
TYPE: DNA
ORGANISM: Drosophila melanogaster
US-09-270-767-1995

Query Match 5.3%; Score 165.8; DB 3; Length 576;

Best Local Similarity 61.3%; Pred. No. 4.7e-34; Matches 285; Conservative 0; Mismatches 177; Indels 3; Gaps 1;

```
Qy 598 CACCATTAACCACTTGGACAACACAGAGATTCAGATTTCCAGCTCTAGAGATTGCC 657
Db 6 CACTTAAGAGATGTGTGTCATCCCGCGGAGCGGACCTTAACCTTTGAGAGCGGT 65
Qy 658 CGTCGCTAGAGATGTATGGAATCCGGTTCACCCGCGCAAGCAGGAGGACGAG 717
Db 66 CGAGAGGTGAGCTGTATGCAATGAATAATCATCCGCGAAGATGTGAAGGGGTCCG 125
Qy 718 ATCAATCTGGCGGTGCAACAGGGAATCTAGTGTTCAGGGTTTCACTAAGATCAAT 777
Db 126 CTTAACTGTGCTGTGGCCACATGGGATCAAGCTTTCCAGAACATCAGCGGATCAAC 185
Qy 778 GCCTTCAACTGGGCAAGGTGCGGAGCTGAGCTTCAAGAGAGCGCTTTCATCAAG 837
Db 186 ACCTTCTCGTGGGCTAAGATACGCAAGATTTCTTCAAGGCAAGGATTTCTGGTCAA 245
Qy 838 CTCGCGGCAAGTCCATAGTGCATCAGAGATACCTTGAATTTCTGATGGCAGTCCG 897
Db 246 CTCGATCCGAGGAGATATGATATTACAGAGATACCTGTGAGTCTTTCGAGGGTCCG 305
Qy 898 GATTTGCAAGTCTTCTGAAAATCTGTGGAATCATCATGCTCTTTAGACTTTT 957
Db 306 AAGAGTGCAGAAATCTTGTGAAAATGCGTGCAGAAATACGAGATTTCTTCCAGTCACT 365
Qy 958 GAAGAGCCCAA---ACCAAGCCCAAGCCGCTCTCTTTAGCCGGGGTCAATTTCCG 1014
Db 366 GCGGTACAAATAATGCGCCAGGCGCAAAATCGGGTCTCTCGGGGTATGATTCCTCCG 425
```

```
Qy 1015 TTCAAGTGTGCACTCAGAAAGCAAGTTCGACTATGTTAAAGA 1059
Db 426 TATAGCGGAAAAACCAAGAAAGATTAATGAGTTGTCGCA 470
```

RESULT 8
US-09-270-767-17277

Sequence 17277, Application US/09270767
Patent No. 6703491
GENERAL INFORMATION:
APPLICANT: Homburger et al.
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
FILE REFERENCE: File Reference: 7326-094
CURRENT APPLICATION NUMBER: US/09/270.767
NUMBER OF SEQ ID NOS: 62517
SOFTWARE: Patent Ver. 2.0
SEQ ID NO 17277
LENGTH: 576
TYPE: DNA
ORGANISM: Drosophila melanogaster
US-09-270-767-17277

Query Match 5.3%; Score 165.8; DB 3; Length 576;
Best Local Similarity 61.3%; Pred. No. 4.7e-34; Matches 285; Conservative 0; Mismatches 177; Indels 3; Gaps 1;

```
Qy 598 CACCATTAACCACTTGGACAACACAGAGATTCAGATTTCCAGCTCTAGAGATTGCC 657
Db 6 CACTTAAGAGATGTGTGTCATCCCGCGGAGCGGACCTTAACCTTTGAGAGCGGT 65
Qy 658 CGTCGCTAGAGATGTATGGAATCCGGTTCACCCGCGCAAGCAGGAGGACGAG 717
Db 66 CGAGAGGTGAGCTGTATGCAATGAATAATCATCCGCGAAGATGTGAAGGGGTCCG 125
Qy 718 ATCAATCTGGCGGTGCAACAGGGAATCTAGTGTTCAGGGTTTCACTAAGATCAAT 777
Db 126 CTTAACTGTGCTGTGGCCACATGGGATCAAGCTTTCCAGAACATCAGCGGATCAAC 185
Qy 778 GCCTTCAACTGGGCAAGGTGCGGAGCTGAGCTTCAAGAGAGCGCTTTCATCAAG 837
Db 186 ACCTTCTCGTGGGCTAAGATACGCAAGATTTCTTCAAGGCAAGGATTTCTGGTCAA 245
Qy 838 CTCGCGCAGATGCCAATAGTGCATCAGAGATACCTTGAATTTCTGATGGCAGTCCG 897
Db 246 CTCGATCCGAGGAGATATGATATTACAGAGATACCTGTGAGTCTTTCGAGGGTCCG 305
Qy 898 GATTTGCAAGTCTTCTGAAAATCTGTGGAATCATCATGCTCTTTAGACTTTT 957
Db 306 AAGAGTGCAGAAATCTTGTGAAAATGCGTGCAGAAATACGAGATTTCTTCCAGTCACT 365
Qy 958 GAAGAGCCCAA---ACCAAGCCCAAGCCGCTCTCTTTAGCCGGGGTCAATTTCCG 1014
Db 366 GCGGTACAAATAATGCGCCAGGCGCAAAATCGGGTCTCTCGGGGTATGATTCCTCCG 425
Qy 1015 TTCAAGTGTGCACTCAGAAAGCAAGTTCGACTATGTTAAAGA 1059
Db 426 TATAGCGGAAAAACCAAGAAAGATTAATGAGTTGTCGCA 470
```

RESULT 9

US-09-848-294-1
Sequence 1, Application US/09848294
Patent No. 6479640
GENERAL INFORMATION:
APPLICANT: Tonks, Nicholas K.
TITLE OF INVENTION: Isolation of A cDNA Encoding A No. 6479640e1
TITLE OF INVENTION: Protein Tyrosine Phosphatase which Localizes to Focal
FILE REFERENCE: CSHL90-04FZA
CURRENT APPLICATION NUMBER: US/09/848,294
CURRENT FILING DATE: 2001-05-03


```
Db 540 AGAGCCCTGCGCCAGGGGGAACTCTCCATCTGAATAAGGAAAGTGCTGGAATGATG 599
Qy 677 GAATCCGGTTGCAACCCGCGCAAGACAGGAGAGCAAGATCAATCTGGCCGTGCA 736
Db 600 GGGTAGACATGCACTGTTGAGGGGAAAGATGCTGTGAATATCTCTTGGACAGACC 659
Qy 737 ACAGGGGAATTTAGTGTTCAGGGTTTCACTAAGATCAATGCTTCAACTGAGCCAAAG 796
Db 660 CGACAGGCAATTAATCTTTGAAGAGCTTAACAAATAGGCTTAATCTTTGACCTTAAA 719
Qy 797 TCGGAGAGCTGACCTTCAGAGAGAGCGCTTCTCATCAAGCTCCGCGCAGATGCCAATA 856
Db 720 TTACCAAAATGATTTTAAABAAGAGCAATTAACACTGCGGTGCTGAGAGATGATGATC 779
Qy 857 -----GTGCGTACCAAGATACCTTGGAAATTCGTATGCGCACTCGGAAATTTTCGAAT 910
Db 780 AGGAGCGTGGCAAGACACGTTGTGTCCGTTTGAACAGTGCAGAGACCTGCAAAAC 839
Qy 911 CCTTCTGGAATAATCTGTGTGAACATCATGCTCTTTTGAAGAGCCCAAAAC 970
Db 840 ACCTTTGAAAGTGTGACGTTGAGACACGCAATTCCTCGACTGCGAGCCGCAAGAAACA 899
Qy 971 CAAAGCCCAA---GCCGTCCTCTTTAGCCGCGGAGTCATCATTTCCGTTCACTGTCGGA 1027
Db 900 GGAATTCGAATAGATCGAATTTATCATGAGCTGCGCTCTGCTTCAAGATTCAGTGGGCGGA 959
Qy 1028 CTCGAGAGCAGGTT 1041
Db 960 CAGATATCAAGCT 973
```

RESULT 11

```
US-09-949-016-3593
; Sequence 3593, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: PseSeq for Windows Version 4.0
; SEQ ID NO 3593
; LENGTH: 2156
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-3593
```

```
Query Match 4.5%; Score 142.4; DB 3; Length 2156;
Best Local Similarity 50.7%; Pred. No. 1.6e-27;
Matches 422; Conservative 0; Mismatches 401; Indels 9; Gaps 3;
```

```
Qy 126 CAAATTCGATGCTGTGATGACACCAAGAGGATTTGAAGTTCACAAAGACTCTG 185
Db 110 CAGCATCTACTTTTGAATGGCGTGTACAGACCTTTAAAGTTACTTAAACAGACACTG 169
Qy 186 GAGGTGCTGCTGATGATGCACTTGAACCACTCAACTGCTGGAAGGTGACTATTTTG 245
Db 170 CAGGTTCTTCTGTGATGTGTGCAACACGCTGGGGTGTACTGAAAAGAAATTTTGG 229
Qy 246 CCTGAGTTTCTGTATCAAAAAAGATCAGGTTG---TGCTGATCTCTTAAACCCAT 302
Db 230 TTATGAGCATGATGACGATCCCTGTGACTCTCTAGATGCTGTGAAGCAAGAACCAT 289
```

```
Qy 303 TGTAAACAGATTAGAGGCCAAAGACGCTGTGTGTTAGATTGTGTGAATTCCTCC 362
Db 290 CAGAGACAGTTTAAAGAGGTTTCCCTGTATCCCTGATCTTTCAGATGATTTTAT 349
Qy 363 GCCTGACACACAACTCCAGAGAACTCAAGAGTACTGTGTGCGCTGAGGTGAA 422
Db 350 ACCTGATCCCAACACTGACAGAAACAAACAGAGCACTGTATTTCTTCAACTGAA 409
Qy 423 GCAGACTTGGCTCAAGGAGTTGACGTGTATGACACAGCGAGCTCTGTGATTTG 482
Db 410 GATGATATTTTGGAGGAGGTTAACCTGCCCCCTTAACTGACGAGTGTTTACGCTC 469
Qy 483 ACATATTGCAATCTGATTTGGGATTTTGTATGAAGCTT---GACAGAGACACTT 539
Db 470 CTATGCGGTACATCTCATTTTGGAGACTATATTTCTTCAATCATCATCCAGCTATCT 529
Qy 540 AGCAAAAATTAATACATACCTCAGCAGACGCACTAGAGACAAATGCTGGAATTTGA 599
Db 530 TTCCGATGACTCATTTATACCGATCAAAATGAGACTTTTAAACAAAGTCGATCTCT 589
Qy 600 CCAAT---AACCATTTGACAAACACACAGACATGATATTTCCAGCTCTAGAGATTGC 656
Db 590 GCATGACAGACAGTGGGCTTAAACAAATCAGAAAGCAATCTGCTATATCAACTAGC 649
Qy 657 CCGTGGCTAGAGATGTATGGAATCCGTTGCAACCCGCGCAAGACAGGAAAGCAGAA 716
Db 650 GCGGACCTTCGACTTTATGAGAGTGAATCTGCAAGTGTAGGATCTGCAATTTAGA 709
Qy 717 GATCAATCTGCGCGTTGCCAAACAGGGAATCTAGTGTTCAGGGTTTCACTAAGATCAA 776
Db 710 CTAATGATTTGGAATTTGCTTCGCGGGGTGTGCTGTGTACCGAAATATCATTTGCAAG 769
Qy 777 TGCCTTCACTGGGCCCAAGAGTGGCAAGCTGAGCTTCAAGAGAGCGCTTCTCATCA 836
Db 770 TTTCTATCTTGGGTGAACATTTCTCAAAATTTCTTCAAAAGAAAGTTCTTCTTACA 829
Qy 837 GCTCGGCGCAGATGCAATATGTCGTAACCAAGATACCTTGAATTCGTATGAGCCAGTGC 896
Db 830 TACGACAGCAAGACAGCTGAATTCAGGGAACATATTTGGCCCTTCAACATGTAATTA 889
Qy 897 GGAATTTCTGCAAGTCTTCTGGAATACTGTGTGGAATCATAGCTTCTT 948
Db 890 CCGATTTGCAAAAATTTGTGGAATCTGTGTGTGACACCATAGCTTCTT 941
```

RESULT 12

```
US-09-799-451-201
; Sequence 201, Application US/09799451
; Patent No. 6783969
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Zhou, Ping
; APPLICANT: Goodrich, Ryle
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyuan
; APPLICANT: Zhang, Jie
; APPLICANT: Xue, Aidong J.
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Ma, Yunding
; APPLICANT: Yamazaki, Victoria
; APPLICANT: Chen, Rui-Hong
; APPLICANT: Wang, Zhiwei
; APPLICANT: Wang, Dunrui
; APPLICANT: Yang, Yonghong
; APPLICANT: Wehrman, Tom
; APPLICANT: Ghosh, Reena
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6783969el Nucleic Acids and
; FILE REFERENCE: 803 Polypeptides
; CURRENT APPLICATION NUMBER: US/09/799,451
; CURRENT FILING DATE: 2001-03-05
```

NUMBER OF SEQ ID NOS: 948
SOFTWARE: PL_FL_genes Version 2.0
SEQ ID NO 201
LENGTH: 3398
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (174)..(2234)
US-09-799-451-201

Query Match 4.3%; Score 133.6; DB 3; Length 3398;
Best Local Similarity 49.5%; Pred. No. 4.7e-25;
Matches 495; Conservative 0; Mismatches 484; Indels 21; Gaps 5;

QY 137 TCGTGTATGACACCCAGAGGCGCTTTGAAGTCCAAAGAGAGTCTCTGGAGAGGCTGCG 196
DB 319 TTCTGTGATGTACTGTAGTGTAGTGTGACCTTGCAGAGAGAGAGAGAGAGAGTGT 378
QY 197 TGGATGACGTTTGCAACCACTCAACCTCGTGGAGAGTACTATTTTGGCTCTGAGTTTC 256
DB 379 TTGATGAGATTATGTACCACTGACCTGACCTGATTAAGAGCACTATTTTGGTCTGAGTTTA 438
QY 257 CTGATCACAAAAAGATCACGGTGTGCTGATCTCTTAAACCCATTTGAAACAGATTAA 316
DB 439 TGGATTACAGACAAAGTAGACATTTGGTGGATGTACAAAAGAGCATCAAAAAGCAAGTAA 498
QY 317 GAAGGCCAAAGACCGTTGTTTAAGTTTGCTGTAATTTCTTCCGCTGACCAACAC 376
DB 499 AATTTGGTTACCCCTATGTCTGATCTTGAAGTTAAGTTTATTTCTTGAAGAGCAATA 558
QY 377 AACTCCAAAGAGACTACAAAGGTACTGCTGCGCTGACAGTGAAGAGAGAGAGAGTGTG 436
DB 559 ACCTTGTGAGAGAGCTAACCCGCTATTTATTTCTTCAAGTTTAAACAGATATTTCTCA 618
QY 437 AAGCAGGTTGACGTGTAAATGACACAGCGGAGCTCTTTGATTTCAACATTTGCAAT 496
DB 619 GTGGAAATTTAGACTGCTCTTGTATATACAGAGTGAATGGAGAGCTTAACTTCAAG 678
QY 497 CTGAGATTGGGATTTTGTATGAAAGCTTGGACG---AAGACCTTAGCAAAAATAAT 553
DB 679 CTGAACTTGGTGTATGTATGATCTTGTGAGCATGTCTTGAACCTTGTCTGAGATTCA 738
QY 554 ACATAC---TCAGCAAGAGCGCACTAGAGAGCAAAATCTGTAATTTCAACATAACACA 610
DB 739 TCGTGTCTATTCAGATGAGAGATGAGATCTGCTATTTTGAAGATGAGAGATATA 798
QY 611 TTGGACAAACACCAAGAGATAGATTTCAAGTCTCTAGAGATTGCGCTCGCTAGAGA 670
DB 799 GAGGTCAAAACACCAAGAGCTGAACCAATATCTGAATAAAGCCAAATGCTAGAAA 858
QY 671 TGTATGGAATCCGGTGTGACCCGGGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 730
DB 859 TGTATGGGGTGTATGATGATGTGTGAAGGCTAGAGATGAGATGCTATGATTTGGAGC 918
QY 731 TTGCAACACAGGAGATTTAGTGTTCAGGGTTTCACTAAGATCAATGCTTCAACTGAG 790
DB 919 TAAACACCAAGAGAGCTGTGTTTGAAGAGAGATACAAAATGGCTATATTTTGGC 978
QY 791 CCAAGGTGCGAAGCTGAGCTTCAAGAGAGAGCGCTTCTCATCAAGCTCGGCGCAGT 850
DB 979 CGAAGATTAACAGATTGATTTTAAAGAGATTAATTAACCTTGGTGTGTGAGAGATG 1038
QY 851 CCAAT-----AGTGGTACCAAGAGATACCTTGAATTCGAGAGGCAATCGGAGATTCT 904
DB 1039 ATGATCAGGGCAAGAT 1098
QY 905 GGAAGTCTTTGGAAGATCTGTGTGAACATCATGCTCTTTAGACTTTTGAAGAGC 964
DB 1099 GGAAGATTTATGGAATGTGTGTGAGAGATCATGCTTTCTTCCGCTTCAAGGCGCG 1158
QY 965 CCAAGCAAGCCCAAGC---CCGTCTTTTACCGGGGGGTATCATTTTGGTTCAATG 1021

DB 1159 TCCAAAGAGTTCTCATGATCATGAGATTATTCGACTAGATCACAGATTATAGTAGT 1218
QY 1022 GTGCGACTCAGAGAGAGGTTCTCGACTATGTTAAAGAGAGAGACATAGAGTGCACT 1081
DB 1219 GGAAGACAGATATGAGACACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATCA 1272
QY 1082 TTGAAGAGAGAGAGAGAGAGATTTCTATTCGGAGCT 1121
DB 1273 TTGAAGAGAGAGAGAGAGAGATTTCTATGAGAGAGAGCT 1312

RESULT 13
US-09-270-767-1542
Sequence 1542; Application US/09270767
Patent No. 6703491

GENERAL INFORMATION:
APPLICANT: Homburger et al.
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
FILE REFERENCE: File Reference: 7326-094
CURRENT APPLICATION NUMBER: US/09/270,767
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 62517
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1542
LENGTH: 449
TYPE: DNA
ORGANISM: Drosophila melanogaster

US-09-270-767-1542

Query Match 3.4%; Score 106.4; DB 3; Length 449;
Best Local Similarity 65.7%; Pred. No. 3.3e-18;
Matches 155; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

QY 2713 CCGACCCGCGCAACAATGAGTGTGACAGTGTGCGACCCGCAACACAGAGCTTCATG 2772
DB 199 CATCCCAAGAGAAATGAGGCTCTGACGCTGTGTGCAATGTGTGCGCACCTTAGACTG 258
QY 2773 GTGACCTTCAGCATGCGAGTGAAGATCATGTTCTGGAACCTGCTGAGAAATTCAA 2832
DB 259 GCGCATACCTGATATCCGCGAGACAGATTGTGCGGTATCTGCTGGAAGTTCAAG 318
QY 2833 AACAGCAACGGGTGCGAAGAGTGTGTTGTTTCAAACTTCTGCTTCTTTAC 2892
DB 319 AATACCTCGGTGCGCAAAAGCTCTGGGTGTGTGCTGCTTGTGTCTGTATCTTAA 378
QY 2893 AATACACACGAGAAATCATCCCTTGCACGCTGCTGCTGCTGCTGCTGCTGCTGCT 2948
DB 379 AATACCTACGAG 434

RESULT 14

US-09-270-767-16824
Sequence 16824; Application US/09270767
Patent No. 6703491

GENERAL INFORMATION:
APPLICANT: Homburger et al.
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
FILE REFERENCE: File Reference: 7326-094
CURRENT APPLICATION NUMBER: US/09/270,767
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 62517
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 16824
LENGTH: 449
TYPE: DNA
ORGANISM: Drosophila melanogaster

US-09-270-767-16824

Query Match 3.4%; Score 106.4; DB 3; Length 449;
Best Local Similarity 65.7%; Pred. No. 3.3e-18;
Matches 155; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

QY 2713 CCGACCCGCGCAACAATGAGTGTGACAGTGTGCGACCCGCAACACAGAGCTTCATG 2772


```
Db      199  CCATCCCGAAGCAATACGGCTCTGCAAGTCTGTTGGCATGTGGCGGCAACGTAGACTG 258
Qy      2773 GTGGACTTCAGCATCCGACGTGAGAAATCAATTGTCTGGAAACCTGCTAGAGAAATTCAA 2832
Db      259  GCGGATCACCCTGATATCCGCGAGACACAGTTGTCGGTTATCTGCTGCGAAAGTTCAAG 318
Qy      2833 AACAGCAACGGGTGGGAGAGAGTGTGGGTGTGTTGACAACTTCGCTGTTCTTCTAC 2892
Db      319  AATAGCTCCGGCTGGCAAAAGCTCTGGGTGTGTTGACGTCCTTTGTCTGTACTTTAC 378
Qy      2893 AATACACACGAGCAATCATCCCTTGCCAGCCTGCTGCTCGGCTACTGCT 2948
Db      379  AAAAGCTACAGAGCAGATTGCACTGCTAGTCTGCGCTATTGGGTTACAGGT 434
```

RESULT 15

```
US-09-621-976-15728
; Sequence 15728, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Joberc, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTe and Encoded Human Proteins.
; FILE REFERENCE: GENSET 054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 15728
; LENGTH: 184
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-15728
```

```
Query Match      3.4%; Score 105.2; DB 3; Length 184;
Best Local Similarity 99.1%; Pred. No. 4.4e-18;
Matches 105; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1692 GTGGTTTCAGAGCAGTGAAGCAAGAGACGCCATGCCGAGACAGTGAAGTCTCAT 1751
Db      79  GTGGTTTCAGAGCAGTGAAGCAAGAGACGCCATGCCGAGACAGTGAAGTCTCAT 138
Qy      1752 ATTCCGAAATTTGAACCTTTGCAAAATTTCTACTAATTTTCTC 1797
Db      139 ATTCCGAAATTTGADCTTTGCAAAATTTCTACTAATTTTCTC 184
```

Search completed: May 22, 2006, 09:18:20
Job time : 543.44 sec

This Page Blank (uspto)

Qy	181	CCCTGGGAAGTGTCTGTGTGATGTCACTGTTTGCAACCACTTCAACCTCTGTGTGAAGGTGACTAT	240
Db	229	CCCTGGGAAGTGTCTGTGTGATGTCACTGTTTGCAACCACTTCAACCTCTGTGTGAAGGTGACTAT	288
Qy	241	TTTGGCTGTGAGTTTCTGTATCACAAAAAAGATCAAGTGTGGCTGTGATCTCTCTAAAAACC	300
Db	289	TTTGGCTGTGAGTTTCTGTATCACAAAAAAGATCAAGTGTGGCTGTGATCTCTCTAAAAACC	348
Qy	301	ATTGTGAAAACAGATTAGAAAGCCAAAGCACGTTGTGTAAAGTTGTGTGAAAATTCCTT	360
Db	349	ATTGTGAAAACAGATTAGAAAGCCAAAGCACGTTGTGTAAAGTTGTGTGAAAATTCCTT	408
Qy	361	CCGCTGTGACACACACACTCCAAAGAACTCAACAAGTACCTGTTTGGCTGTCAAGTG	420
Db	409	CCGCTGTGACACACACACTCCAAAGAACTCAACAAGTACCTGTTTGGCTGTCAAGTG	468
Qy	421	AAGCAGGAATTGGCTCAAGGGCAGGTGTGACGTATATGACACAGGAGCAGCTCTCTGATT	480
Db	469	AAGCAGGAATTGGCTCAAGGGCAGGTGTGACGTATATGACACAGGAGCAGCTCTCTGATT	528
Qy	481	TCACACATTGTGCAATCTGAGATTGGGGATTTTGTATGAAAGCTTGTGACAGAGACACTTA	540
Db	529	TCACACATTGTGCAATCTGAGATTGGGGATTTTGTATGAAAGCTTGTGACAGAGACACTTA	588
Qy	541	GCAAAAAATTAATACATACCTTACGACAGACGCACTAGAGACAAATGTGTGAATTTTAC	600
Db	589	GCAAAAAATTAATACATACCTTACGACAGACGCACTAGAGACAAATGTGTGAATTTTAC	648
Qy	601	CAATAACCAATTGGACAAACACACAGACAGATTCAGAGTTCAGAGTCTAGAGATTGGCCGT	660
Db	649	CAATAACCAATTGGACAAACACACAGACAGATTCAGAGTTCAGAGTCTAGAGATTGGCCGT	708
Qy	661	CGGCTAGAGATGTATGGAATCCGGTTGCAACCGGCAAGGACAGGGAAGGACAGAGATC	720
Db	709	CGGCTAGAGATGTATGGAATCCGGTTGCAACCGGCAAGGACAGGGAAGGACAGAGATC	768
Qy	721	AATCTGGCCGTTGCCAACAACGGGAATTTCTAGTGTTCACGGTTTCACTAAGATCAATGCC	780
Db	769	AATCTGGCCGTTGCCAACAACGGGAATTTCTAGTGTTCACGGTTTCACTAAGATCAATGCC	828
Qy	781	TTCAACTGGGCCAAGTGTGCGGAAGCTGAGCTTCAGAGAAAGCGTTTCTGATCAAGTCTC	840
Db	829	TTCAACTGGGCCAAGTGTGCGGAAGCTGAGCTTCAGAGAAAGCGTTTCTGATCAAGTCTC	888
Qy	841	CGGCGAGATGCCAATATGTGCTGACAGAGTACCTTGGAAATTCGTATGGCAGTCCGAGAT	900
Db	889	CGGCGAGATGCCAATATGTGCTGACAGAGTACCTTGGAAATTCGTATGGCAGTCCGAGAT	948
Qy	901	TTCTGTGCAAGTCTTCTGTGAAAAATCTGTGTGAACATACGCTTCTTTAGACTTTTGA	960
Db	949	TTCTGTGCAAGTCTTCTGTGAAAAATCTGTGTGAACATACGCTTCTTTAGACTTTTGA	1008
Qy	961	GAGCCCAACCAAAAGCCCAAGCCGCTCTTTACCGGGGTCAATCTTCGGTCAAT	1020
Db	1009	GAGCCCAACCAAAAGCCCAAGCCGCTCTTTACCGGGGTCAATCTTCGGTCAAT	1068
Qy	1021	GGTGCGACTCAAGAGAGGTTTCTGACTATGTTTAAAGAAAGAGACATAAGAGGTGACG	1080
Db	1069	GGTGCGACTCAAGAGAGGTTTCTGACTATGTTTAAAGAAAGAGACATAAGAGGTGACG	1128
Qy	1081	TTTGAAGGAGACACAGCAAGATTCATCTATCCGAGCCCTTGCTTCAAGCTTACAGAA	1140
Db	1129	TTTGAAGGAGACACAGCAAGATTCATCTATCCGAGCCCTTGCTTCAAGCTTACAGAA	1188
Qy	1141	CTGAATTCGGAATGTGTGAGAGCTCTACAGCAAGACACAGACTTTTACATTTTGGAAAGGT	1200
Db	1189	CTGAATTCGGAATGTGTGAGAGCTCTACAGCAAGACACAGACTTTTACATTTTGGAAAGGT	1248
Qy	1201	GCGGAATCTCAAGGGGGCCAGAGCTGCGGCGCAGAAAGAAACGAAAGTATTTCGCGCGGG	1260
Db	1249	GCGGAATCTCAAGGGGGCCAGAGCTGCGGCGCAGAAAGAAACGAAAGTATTTCGCGCGGG	1308
Qy	1261	GAGCCGGGTGCAACCCGAGCCCTGTGCGCGAGAGAAAGCCCGCGGTATACAAAGCAAGCG	1320

Dp	1309	GAGCGGGGGTGCACCCGAGCCTTGCCCGAGAGAGGCCCGGGTTAACAGCAGCGC	1368
Qy	1321	GACCGAGCCGCTTCGCGGCCCAACGAGAGAGGTCTTTAAGATAGAGCCAG	1380
Dp	1369	GACGAGCCGCGCTCGCGGCCCAACGAGAGAGAGGTCTTTAAGATAGAGCCAG	1428
Qy	1381	CAGAGTAACTCAGCCCCCGCAGCCCAAGCAGAGCTCCTGACTGGCACTCTCACCTT	1440
Dp	1429	CAGAGTAACTCAGCCCCCGCAGCCCAAGCAGAGCTCCTGACTGGCACTCTCACCTT	1488
Qy	1441	TCGAGCGTCTGTGAACCTCGCAGGGGGGAGTGGCCCTGTCACAACGATCCTTGTCCTCC	1500
Dp	1489	TCGAGCGTCTGTGAACCTCGCAGGGGGGAGTGGCCCTGTCACAACGATCCTTGTCCTCC	1548
Qy	1501	AACCTGAGCCCGCAGCAACCAAGCAGGCTCTCCTTGATCAGCCCGCTGTAATGACAG	1560
Dp	1549	AACCTGAGCCCGCAGCAACCAAGCAGGCTCTCCTTGATCAGCCCGCTGTAATGACAG	1608
Qy	1561	GCCCTGCCCCCGGAGCGAGTAGAGATGAGGGCCGAGAGAAAGATTCCTCACTGATATA	1620
Dp	1609	GCCCTGCCCCCGGAGCGAGTAGAGATGAGGGCCGAGAGAAAGATTCCTCAACTGATATA	1668
Qy	1621	GCGACTTCAATAGTAGAGAGGTCTACACCGAGGGAACTATCTGAAGAGATCTTCGA	1680
Dp	1669	GCGACTTCAATAGTAGAGAGGTCTACACCGAGGGAACTATCTGAAGAGATCTTCGA	1728
Qy	1681	GTTATCACTTCGTGGTTTCAGAGCAGTAGAGCAAAAGAGCGCATGCCGAGAGCAGCTG	1740
Dp	1729	GTTATCACTTCGTGGTTTCAGAGCAGTAGAGCAAAAGAGCGCATGCCGAGAGCAGCTG	1788
Qy	1741	AAAGTCTCATATTCGCCGAATTTTGAACCTTTGCACAAAATTTCACTAAATTTTCTCAG	1800
Dp	1789	AAAGTCTCATATTCGCCGAATTTTGAACCTTTGCACAAAATTTCACTAAATTTTCTCAG	1848
Qy	1801	GAAATTGAGCAACGACTTGGCCCTGTGGGAGGGCCGCTCAATGCCCAATCCAGAGATTAC	1860
Dp	1849	GAAATTGAGCAACGACTTGGCCCTGTGGGAGGGCCGCTCAATGCCCAATCCAGAGATTAC	1908
Qy	1861	CAAGAATTCGCGCATGTCATGCTGTAAGAAACATTCAAGGGCATGAAAGCAGCTGCGGCTCAC	1920
Dp	1909	CAAGAATTCGCGCATGTCATGCTGTAAGAAACATTCAAGGGCATGAAAGCAGCTGCGGCTCAC	1968
Qy	1921	CTGTGGAAGCACAGCGAGGCCCTTGGAGGCCCTTGGAGATGAAATCAAGAGCTCCCGCGG	1980
Dp	1969	CTGTGGAAGCACAGCGAGGCCCTTGGAGGCCCTTGGAGATGAAATCAAGAGCTCCCGCGG	2028
Qy	1981	CTGGAAGAACTTCTGCAAGACCTTTGAGCTCAGAAAGGTGTGTAACTCAACGGCTCAACCC	2040
Dp	2029	CTGGAAGAACTTCTGCAAGACCTTTGAGCTCAGAAAGGTGTGTAACTCAACGGCTCAACCC	2088
Qy	2041	TTTCTCTCGCGGCACATGCAACCGGCTCATGCACTAACAGAGAGCTTGAGACGGCTGTGC	2100
Dp	2089	TTTCTCTCGCGGCACATGCAACCGGCTCATGCACTAACAGAGAGCTTGAGACGGCTGTGC	2148
Qy	2101	AAACACCAACCCCGCGAGCCACGCGCAGCTTCAGGAGCTGCCAGCCGCTTTGGCAGAGATC	2160
Dp	2149	AAACACCAACCCCGCGAGCCACGCGCAGCTTCAGGAGCTGCCAGCCGCTTTGGCAGAGATC	2208
Qy	2161	ACGAGATGTGTGACAGCTCTCCACGATGATGATCAAGATGAGAAATTTTCCAGAACTGTG	2220
Dp	2209	ACGAGATGTGTGACAGCTCTCCACGATGATGATCAAGATGAGAAATTTTCCAGAACTGTG	2268
Qy	2221	CACGAACTCAAGAAAGATTTGATTGGGATTTGACAAATTTTGTTGCTCGGGAAGGAGATTCC	2280
Dp	2269	CACGAACTCAAGAAAGATTTGATTGGGATTTGACAAATTTTGTTGCTCGGGAAGGAGATTCC	2328
Qy	2281	ATTCGCTGTGGCAGACCTCAGCAAGCTCTCGGGAGAGGGGCTCCAGACGCGATGTTCTTC	2340
Dp	2329	ATTCGCTGTGGCAGACCTCAGCAAGCTCTCGGGAGAGGGGCTCCAGACGCGATGTTCTTC	2388
Qy	2341	CTGTTCAACGAGTCTGTGCTATACAGAGCCGGGGGCTGACGCGCTCAATCACTTTAAA	2400

Db 2389 CTGTTCAACGACGCTCTGCTATATACAGAGCCGGGGCTGACGGCTCTCAATCAGTTTAA 2448
QY 2401 GTCCACGGGACAGTCCCGCTCTTATGGCATATGAGAGGCGAAGCGAGTGGGG 2460
Db 2449 GTCCACGGGACAGTCCCGCTCTTATGGCATATGAGAGGCGAAGCGAGTGGGG 2508
QY 2461 GTGCCCCACTGCTGACCCCTCCGGGGCGAGGCGAGTCCATCATCTGTCGGCCAGTTCT 2520
Db 2509 GTGCCCCACTGCTGACCCCTCCGGGGCGAGGCGAGTCCATCATCTGTCGGCCAGTTCT 2568
QY 2521 CGGTCCGAGATGAGAGAGTGGGTTGAGAGCATTCAGATGCCATTGACCTGGCGAGAG 2580
Db 2569 CGGTCCGAGATGAGAGAGTGGGTTGAGAGCATTCAGATGCCATTGACCTGGCGAGAG 2628
QY 2581 AGCAGAGCCCGCCCTCTGAGTTCCTGGCCAGAGCCCTCTCAACAAGTCCCTGAT 2640
Db 2629 AGCAGAGCCCGCCCTCTGAGTTCCTGGCCAGAGCCCTCTCAACAAGTCCCTGAT 2688
QY 2641 GAAGCCAGCCGCTGACAGAGATGAGAGATGAGAGCCCTCTGCGCAGATGCTG 2700
Db 2689 GAAGCCAGCCGCTGACAGAGATGAGAGATGAGAGCCCTCTGCGCAGATGCTG 2748
QY 2701 GAGCCGACGAGCCCGCAGACCGCGGCAACACATGCTGCACTGTGCGACCGCAACCC 2760
Db 2749 GAGCCGACGAGCCCGCAGACCGCGGCAACACATGCTGCACTGTGCGACCGCAACCC 2808
QY 2761 AGGCTTCCATGGTGAATTCACAGATCGAGTGGAGATTCAGTTGCTGGAACTGCTG 2820
Db 2809 AGGCTTCCATGGTGAATTCACAGATCGAGTGGAGATTCAGTTGCTGGAACTGCTG 2868
QY 2821 AGGAAATTCAAAACAGCAACGAGGAGAGAGTGGGTTGTTCAAACTTCTGC 2880
Db 2869 AGGAAATTCAAAACAGCAACGAGGAGAGAGTGGGTTGTTCAAACTTCTGC 2928
QY 2881 CTGTTCTTCAAAATCACACAGACAAATCATCCCTTTCAGAGCTCTGCTGCGC 2940
Db 2929 CTGTTCTTCAAAATCACACAGACAAATCATCCCTTTCAGAGCTCTGCTGCGC 2988
QY 2941 TACTGCTCAACATCCCTCTGAGTCCGAGAACATCAAGAAAGATCAAGTTCAGAGCTG 3000
Db 2989 TACTGCTCAACATCCCTCTGAGTCCGAGAACATCAAGAAAGATCAAGTTCAGAGCTG 3048
QY 3001 CACTTCAAGTCCCACTTACTTCAAGGCGGAAAGCGAGTCAAGTTCAGAAAGTGG 3060
Db 3049 CACTTCAAGTCCCACTTACTTCAAGGCGGAAAGCGAGTCAAGTTCAGAAAGTGG 3108
QY 3061 ATGGAAGTATCGAGTCCAGCAGCTGCTGCGAGCCCAAGTTCAGAAAGTGG 3120
Db 3109 ATGGAAGTATCGAGTCCAGCAGCTGCTGCGAGCCCAAGTTCAGAAAGTGG 3168
QY 3121 GAGTCTCTGTGTAT 3135
Db 3169 GAGTCTCTGTGTAT 3183

RESULT 2
US-10-921-707-18
Sequence 18, Application US/10921707
Publication No. US20050003447A1
GENERAL INFORMATION:
APPLICANT: INCYTE PHARMACEUTICALS, INC.
APPLICANT: LAL, Preeti
APPLICANT: TANG, Y. Tom
APPLICANT: YUE, Henry
APPLICANT: HILLMAN, Jennifer L.
APPLICANT: BANDMAN, Olga
APPLICANT: CORLEY, Neil C.
APPLICANT: GUEGLER, Karl J.
APPLICANT: PATTERSON, Chandra
APPLICANT: AZIMZAI, Yaida
APPLICANT: BAUGHN, Mariah R.
TITLE OF INVENTION: HUMAN CYTOSKELETON ASSOCIATED PROTEINS
FILE REFERENCE: PF-0594 PCT

Query Match 99.9%; Score 3131.8; DB 9; Length 4687;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3133; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
CURRENT APPLICATION NUMBER: US/10/921,707
CURRENT FILING DATE: 2004-08-19
PRIOR APPLICATION NUMBER: US/09/786,797
PRIOR FILING DATE: 2001-06-25
PRIOR APPLICATION NUMBER: 09/156,470; Unassigned; 60/131,321
PRIOR FILING DATE: 1998-09-18; 1998-09-18; 1999-04-27
NUMBER OF SEQ ID NOS: 32
SOFTWARE: PERL Program
SEQ ID NO 18
LENGTH: 4687
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No: 1320252
US-10-921-707-18
Query Match 99.9%; Score 3131.8; DB 9; Length 4687;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3133; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
1 ATGGGAAATATGAGCAGAGCCCGACCCAGATTCAGACTGGGGCCCCGGAAATTTG 60
90 ATGGGAAATATGAGCAGAGCCCGACCCAGATTCAGACTGGGGCCCCGGAAATTTG 149
61 GGGATGATACCTTGGGAAAGTGGGACAGAGCCCGGCAACCTTCAGGAAATTCG 120
150 GGGATGATACCTTGGGAAAGTGGGACAGAGCCCGGCAACCTTCAGGAAATTCG 209
121 TCATCAAAATCAGATGCTGATGACACCCAGAGGCAATTTGAAGTTCAACAAGAGCT 180
210 TCATCAAAATCAGATGCTGATGACACCCAGAGGCAATTTGAAGTTCAACAAGAGCT 269
181 CTTGGGAAAGTCTGCTGATGACAGTTTGAACACCTTACCTGTTGAAGTGAAT 240
270 CTTGGGAAAGTCTGCTGATGACAGTTTGAACACCTTACCTGTTGAAGTGAAT 329
241 TTTGGCTGAGTTTCTGATCAAAAAGATCAGGTTGGCTGGATCTCTAAACC 300
330 TTTGGCTGAGTTTCTGATCAAAAAGATCAGGTTGGCTGGATCTCTAAACC 389
301 ATTTGAAACAGATTTAGAGGCAAGCAAGTGTGTTAAGTTGTGTGAAATCTTT 360
390 ATTTGAAACAGATTTAGAGGCAAGCAAGTGTGTTAAGTTGTGTGAAATCTTT 449
361 CCGCTGACCAACCACTCCAGAGAACTCAAGATACCTGTTGGCTGCAAGT 420
450 CCGCTGACCAACCACTCCAGAGAACTCAAGATACCTGTTGGCTGCAAGT 509
421 AGCAGGACTTGGCTCAAGGAGTTGAGTGAATGACAGAGGCAAGCTCTTGAAT 480
510 AGCAGGACTTGGCTCAAGGAGTTGAGTGAATGACAGAGGCAAGCTCTTGAAT 569
481 TCACACATTTGCAATCTGAGATTTGGGAAATTTGATGAAGCTTGGACAGAGCACTTA 540
570 TCACACATTTGCAATCTGAGATTTGGGAAATTTGATGAAGCTTGGACAGAGCACTTA 629
541 GCAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 600
630 GCAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 689
601 CATTAACCAATTTGCAACCAAGAGAAATGAGATTTCACTCTTGAAGTTCGCT 660
690 CATTAACCAATTTGCAACCAAGAGAAATGAGATTTCACTCTTGAAGTTCGCT 749
661 CGGCTAGAGATGATGAATTCGGTTGCAACCGGCAAGAGCAAGGAAAGCAAGATC 720
750 CGGCTAGAGATGATGAATTCGGTTGCAACCGGCAAGAGCAAGGAAAGCAAGATC 809
721 AATCTGGCCGTTGCAACCGGAAATTTCAAGTTTCAAGGTTTCACTTAAGTCAATGCC 780
810 AATCTGGCCGTTGCAACCGGAAATTTCAAGTTTCAAGGTTTCACTTAAGTCAATGCC 869

840
840
929
900
989
960
1049
1020
1109
1080
1169
1140
1229
1200
1289
1260
1349
1320
1409
1380
1469
1440
1529
1500
1589
1560
1649
1620
1709
1680
1769
1740
1829
1800
1889
1860
1949
1920

2009
1980
2069
2040
2129
2100
2189
2160
2249
2220
2309
2280
2369
2340
2429
2400
2489
2460
2549
2520
2609
2669
2729
2700
2789
2849
2820
2880
2969
2940
3029
3000

Db 3030 TACTCGCTACACATCCCTCTGAGTCCGAGNATCCAGAAAGACTAGTGTCAAGCTG 3089
Qy 3001 CACTTCAAGTCCCACTCTACTACTTCAAGGCGGAAAGCGAGTACAGTTTCAAGAGTGG 3060
Db 3090 CACTTCAAGTCCCACTCTACTACTTCAAGGCGGAAAGCGAGTACAGTTTCAAGAGTGG 3149
Qy 3061 ATGAAAGTATCCGAGTGCACCAAGCTGTGCTCGGACCCCACTGTGTGAGCCACAA 3120
Db 3150 ATGAAAGTATCCGAGTGCACCAAGCTGTGCTCGGACCCCACTGTGTGAGTCAAAA 3209
Qy 3121 GAGTCTTGTGTAT 3135
Db 3210 GAGTCTTGTGTAT 3224

RESULT 3
US-10-450-763-564
Sequence 564, Application US/10450763
Publication No. US20050196754A1
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc
TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
FILE REFERENCE: 790CIP3/US
CURRENT APPLICATION NUMBER: US/10/450,763
CURRENT FILING DATE: 2003-06-11
PRIOR APPLICATION NUMBER: PCT/US01/08631
PRIOR FILING DATE: 2001-03-30
PRIOR APPLICATION NUMBER: 09/540,217
PRIOR FILING DATE: 2000-03-31
PRIOR APPLICATION NUMBER: 09/649,167
PRIOR FILING DATE: 2000-08-23
NUMBER OF SEQ ID NOS: 60736
SOFTWARE: Custom
SEQ ID NO 564
LENGTH: 3187
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SIMILAR
LOCATION: (49)..(1371)
OTHER INFORMATION: 92% homologous to Homo sapiens CDBP, accession number
US-10-450-763-564

Query Match 99.5%; Score 3119.2; DB 10; Length 3187;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3132; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

Qy 1 ATGGAGAAATAGAGCAGAGCCGACCCAGATCAAGCTGGGGGCCCCGAAATTCG 60
Db 49 ATGGAGAAATAGAGCAGAGCCGACCCAGATCAAGCTGGGGGCCCCGAAATTCG 108
Qy 61 GGGATCAGTACCTTGGACCTGGACAGAAAGCCGCCCCCAACCTTCAGAAAACTCGTG 120
Db 109 GGGATCAGTACCTTGGACCTGGACAGAAAGCCGCCCCCAACCTTCAGAAAACTCGTG 168
Qy 121 TCATCAAAATCCAGATGCTGATGACACCCAGAGGCAATTTGAAGTTCAAAAAGCT 180
Db 169 TCATCAAAATCCAGATGCTGATGACACCCAGAGGCAATTTGAAGTTCAAAAAGCT 228
Qy 181 CTTGGAGAGGTGCTGCTGATGATGACAGTTTGCACCACTCAACCTCGTGAAGGTAAT 240
Db 229 CTTGGAGAGGTGCTGCTGATGATGACAGTTTGCACCACTCAACCTCGTGAAGGTAAT 288
Qy 241 TTTGGCTTGAAGTTTCTGATCACAAAAGATCAAGGTGTGGTGAATCTCTAAACC 300
Db 289 TTTGGCTTGAAGTTTCTGATCACAAAAGATCAAGGTGTGGTGAATCTCTAAACC 348
Qy 301 ATTGTGAACAGATTGAAGGCCAAGACAGTTGTTTAAGTTTGGTGAATCTCTT 360
Db 349 ATTGTGAACAGATTGAAGGCCAAGACAGTTGTTTAAGTTTGGTGAATCTCTT 408
Qy 361 CGCGCTGACCAACACTCAAGAAAGACTCAAGGTAAGTCTGTTGCGGCTGCAAGTG 420

Db 409 CGCGCTGACCAACAACTTCAAGAAAGACTCAAGAGTACTGTTCCGGCTGCAAGTG 468
Qy 421 AAGCAGACTTGGCTCAAGGCAAGTGAAGTGAATGACACAGCGAGCTCTTGTGAT 480
Db 469 AAGCAGACTTGGCTCAAGGCAAGTGAAGTGAATGACACAGCGAGCTCTTGTGAT 528
Qy 481 TCACACTTGTGAATCTGATTTGGGATTTTGAAGAACCTTGAACAAGAGCACTTA 540
Db 529 TCACACTTGTGAATCTGATTTGGGATTTTGAAGAACCTTGAACAAGAGCACTTA 588
Qy 541 GCAAAAATTAATATACCTCTAGCAAGACGACTAGAGGACAAATCTGTGAATTTAC 600
Db 589 GCAAAAATTAATATACCTCTAGCAAGACGACTAGAGGACAAATCTGTGAATTTAC 648
Qy 601 CATPACCACTTGGACAAACACGACGAAATCCAGTCTTGAAGATTTGCCGT 660
Db 649 CATPACCACTTGGACAAACACGACGAAATCCAGTCTTGAAGATTTGCCGT 708
Qy 661 CGGCTGAGATGATGAAATCCGTTGACCCCGCCAAAGACAGGAAAGCACGAAGATC 720
Db 709 CGGCTGAGATGATGAAATCCGTTGACCCCGCCAAAGACAGGAAAGCACGAAGATC 768
Qy 721 AATCTGCGCTTGCACCAACGAGAAATCTAGTGTTCAGAGGTTTCACTAAGATCAATGCC 780
Db 769 AATCTGCGCTTGCACCAACGAGAAATCTAGTGTTCAGAGGTTTCACTAAGATCAATGCC 828
Qy 781 TTCACTGGGCAAGGTGCGAAGCTGAGCTTCAAGAGAAAGCGTTTCTCATCAAGCTC 840
Db 829 TTCACTGGGCAAGGTGCGAAGCTGAGCTTCAAGAGAAAGCGTTTCTCATCAAGCTC 888
Qy 841 CGGCCAGATGCCAAATAGTGTGTAACAGATACCTTGAATTCCTGATGGCAATCGGAGAT 900
Db 889 CGGCCAGATGCCAAATAGTGTGTAACAGATACCTTGAATTCCTGATGGCAATCGGAGAT 948
Qy 901 TTCTGCAAGCTCTTGGAAAAATCTGTGTGAACATCATCTCTTTAGACTTTTGA 960
Db 949 TTCTGCAAGCTCTTGGAAAAATCTGTGTGAACATCATCTCTTTAGACTTTTGA 1008
Qy 961 GAGCCCAACCAAGCCCAAGCCGCTCTTTAGCCGGGGGTTCATTTCCGTTCAAGT 1020
Db 1009 GAGCCCAACCAAGCCCAAGCCGCTCTTTAGCCGGGGGTTCATTTCCGTTCAAGT 1068
Qy 1021 GGTGGAATCAGAAAGAGTGTCTGACTAATGTTAAAGAGAGAGATTAAGAGTGCAG 1080
Db 1069 GGTGGAATCAGAAAGAGTGTCTGACTAATGTTAAAGAGAGAGATTAAGAGTGCAG 1128
Qy 1081 TTTGAAGAGAGACAGCAAGATTTCAATCTATCCGAGCTTCTTCAAGCTTACAGAA 1140
Db 1129 TTTGAAGAGAGACAGCAAGATTTCAATCTATCCGAGCTTCTTCAAGCTTACAGAA 1188
Qy 1141 CTGAATTCGAAAGTGTGAGAGCTCTCAGACAGACCAAGCTTACATTTGAGAGA-AGG 1199
Db 1189 CTGAATTCGAAAGTGTGAGAGCTCTCAGACAGACCAAGCTTACATTTGAGAGAAGG 1248
Qy 1200 TGGCGAATCTCCAGGGGGCCAGAGCTGCGGCGAAGAAAGAAACGAAAGTTTCCCGCGG 1259
Db 1249 TGGCGAATCTCCAGGGGGCCAGAGCTGCGGCGAAGAAAGAAACGAAAGTTTCCCGCGG 1308
Qy 1260 GGAAGCCGGGGTGCACCCGAGCCCTGCGCGAGAGAAAGCCCGCGGTTAAACAAGAGG 1319
Db 1309 GGAAGCCGGGGTGCACCCGAGCCCTGCGCGAGAGAAAGCCCGCGGTTAAACAAGAGG 1368
Qy 1369 GGAAGCCGGGGTGCACCCGAGCCCTGCGAGAGAAAGAGAGGTGCTTAAGGATGAGACCCA 1428
Db 1380 GGAAGCCGGGGTGCACCCGAGCCCTGCGAGAGAAAGAGAGGTGCTTAAGGATGAGACCCA 1439
Qy 1429 GGAAGCCGGGGTGCACCCGAGCCCTGCGAGAGAAAGAGAGGTGCTTAAGGATGAGACCCA 1488
Db 1440 TTTCCGAGTGTCTGTGAATCTGCAAGGGGGAGTGGCCCTTGCACAGTACCTTGTCTCC 1499

1489 TTCCGAGCTGTCTGTGTAATCCGACGGGGAGTGGCCCTCCGCAACGTGACTCTCC 1548
1500 CAACCTGACCCCGGACACCAAGAGGCTCTCCCTTGATCAAGCCGCTGTGAATGACA 1559
1549 CAACCTGACCCCGGACACCAAGAGGCTCTCCCTTGATCAAGCCGCTGTGAATGACA 1608
1560 GGCCTGCCCCCGGACGAGATGAGGATGAGGCGCGGAGAGAGATTCCTCAACTGATTA 1619
1609 GGCCTGCCCCCGGACGAGATGAGGATGAGGCGCGGAGAGAGATTCCTCAACTGATTA 1668
1620 AGCGTACTTCAATGCTAAGAAAGTGTCTACACCGAGCGCAATATCTGAGATCTCA 1679
1669 ACCGTACTTCAATGCTAAGAAAGTGTCTACACCGAGCGCAATATCTGAGATCTCA 1728
1680 AGTATCACTGTGTGTTCAGAGCAAGTGAAGCAAGAGGAGCGCATGCGGAGCACT 1739
1729 AGTATCACTGTGTGTTCAGAGCAAGTGAAGCAAGAGGAGCGCATGCGGAGCACT 1788
1740 GAAAGTCTCATATTCGCAATTTTGAACCTTTGCAAAATTTTCATATTTTCTCA 1799
1789 GAAAGTCTCATATTCGCAATTTTGAACCTTTGCAAAATTTTCATATTTTCTCA 1848
1800 GGAATTTGACCAACGACTTGGCTGTGGAGAGGCGCTCAATGCCAATACAGATTA 1859
1849 GGAATTTGACCAACGACTTGGCTGTGGAGAGGCGCTCAATGCCAATACAGATTA 1908
1860 CCAAGAAATCGGAGATGTATGTCTGAAGAACATTGAGGAGATGAGCACTGGCGCTCA 1919
1909 CCAAGAAATCGGAGATGTATGTCTGAAGAACATTGAGGAGATGAGCACTGGCGCTCA 1968
1920 CCGTGTGAAGCAACGAGAGGCTTTGAGAGGCGCTGAAGAAATGGAATCAAGAGCTCCGAGCG 1979
1969 CCGTGTGAAGCAACGAGAGGCTTTGAGAGGCGCTGAAGAAATGGAATCAAGAGCTCCGAGCG 2028
1980 GCTGTGAAGCACTTGTGAGAGACTTTGAGGCTGCAAGAGGTGTGTACTTACCGTCAAC 2039
2029 GCTGTGAAGCACTTGTGAGAGACTTTGAGGCTGCAAGAGGTGTGTACTTACCGTCAAC 2088
2040 CTTCCCTCTGCGGCGCACTGCAACGAGCTCATGCACTCAAGACAGTCTCTGAGCGGCTGTG 2099
2089 CTTCCCTCTGCGGCGCACTGCAACGAGCTCATGCACTCAAGACAGTCTCTGAGCGGCTGTG 2148
2100 CAAACACCAACCGGCGGAGCAAGCGGCTTCAAGGAGCTGCGGAGCGCTTTGCGAGAGAT 2159
2149 CAAACACCAACCGGCGGAGCAAGCGGCTTCAAGGAGCTGCGGAGCGCTTTGCGAGAGAT 2208
2160 CACGAGATGTGTGACAGCTCCACGCTACAGATCAAGATGAGAGAAATTTCCAGAAAGCT 2219
2209 CACGAGATGTGTGACAGCTCCACGCTACAGATCAAGATGAGAGAAATTTCCAGAAAGCT 2268
2220 GCACGAACTCAAGAAAGATTTGATTTGAGCAATCTTGTGTGCTCCGAGAAAGGAGATT 2279
2269 GCACGAACTCAAGAAAGATTTGATTTGAGCAATCTTGTGTGCTCCGAGAAAGGAGATT 2328
2280 CATCCGTGTGGGAGGCTTCAAGCAAGCTCTCGGAGAAAGGAGCTTCCAGACGCAATGTTCT 2339
2329 CATCCGTGTGGGAGGCTTCAAGCAAGCTCTCGGAGAAAGGAGCTTCCAGACGCAATGTTCT 2388
2340 CCGTGTCAACGAGCTCTGTCTATACAGAGCCGCGGAGCTGACGCGCTCCAAATCAAGTTTAA 2399
2389 CCGTGTCAACGAGCTCTGTCTATACAGAGCCGCGGAGCTGACGCGCTCCAAATCAAGTTTAA 2448
2400 AGTCCACGGGAGAGCTCCGCTCTATGAGCATGAGATTTGAGAGAGAGGAGAGAGAGAGG 2459
2449 AGTCCACGGGAGAGCTCCGCTCTATGAGCATGAGATTTGAGAGAGAGGAGAGAGAGG 2508
2460 GGTGCCCACTGCTGACCTTCCGGGAGCAAGCGGAGATTCATCATGTGCGCGCACTTTC 2519
2509 GGTGCCCACTGCTGACCTTCCGGGAGCAAGCGGAGATTCATCATGTGCGCGCACTTTC 2568
2520 TCGGTCCGAGATGAGAAAGTGGTTGAGACATCCAGATGCGCATTTGACCTGGCGGAGAA 2579
2569 TCGGTCCGAGATGAGAAAGTGGTTGAGACATCCAGATGCGCATTTGACCTGGCGGAGAA 2628

2580 GAGCAGACGCCCCGCGCTGAGTTCCTGGACAGAGCCCCCTGACAAAGATCCCTGA 2639
2629 GAGCAGACGCCCCGCGCTGAGTTCCTGGACAGAGCCCCCTGACAAAGATCCCTGA 2688
2640 TGAAGCCACCGGAGCTGACAGAGATCAGAGATGACCTGAGCGCTCGGCAATCGCT 2699
2689 TGAAGCCACCGGAGCTGACAGAGATCAGAGATGACCTGAGCGCTCGGCAATCGCT 2748
2700 GGAAGGCGGAGGCGCGGACCGGAGCAACATGTGTGACGAGTGTGAGCAACGAGCAAC 2759
2749 GGAAGGCGGAGGCGCGGACCGGAGCAACATGTGTGACGAGTGTGAGCAACGAGCAAC 2808
2760 CAGCGTCTCAGTGTGACCTTCAAGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 2819
2809 CAGCGTCTCAGTGTGACCTTCAAGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 2868
2820 GAGAAATTTCAAAACAGCAACGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 2879
2869 GAGAAATTTCAAAACAGCAACGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 2928
2880 CCGTGTCTTCAATATCAACAGAGCAATATCCCTTGCAGCTGCTGCTGCTGCTGCTGCTG 2939
2929 CCGTGTCTTCAATATCAACAGAGCAATATCCCTTGCAGCTGCTGCTGCTGCTGCTGCTG 2988
2940 CTAAGTCTCAGTGTGACCTTCAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 2999
2989 CTAAGTCTCAGTGTGACCTTCAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 3048
3000 GCACTTCAAGTGTGACCTTCAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 3059
3049 GCACTTCAAGTGTGACCTTCAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 3108
3060 GATGAGATGATCCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3119
3109 GATGAGATGATCCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3168
3120 AGAGTCTCTTGTGTAT 3135
3169 AGAGTCTCTTGTGTAT 3184

RESULT 4
US-09-925-297-220
; Sequence 220, Application US/09925297
; Patent No. US20020081659A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA105
; CURRENT APPLICATION NUMBER: US/09/925,297
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05989
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; NUMBER OF SEQ ID NOS: 928
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 220
; LENGTH: 3094
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (4)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-297-220

Query Match 45.7%; Score 1431.2; DB 3; Length 3094;
Best Local Similarity 99.8%; Pred. No. 0;

270 GGACAAACACATGAAATATTTGACATGAGCCTAAATGCGATGCGCAGGATTTACTGAC 329
QY
201 TGCAGTTTGCAACCACTCAACCTGCGAGAGTGACTATTTGGCTCGAGTTTCTGA 260
Db
330 ACAAGTGTGAGACGTTTAACTGGTGAATGTGACTTCTGGAGATGAGTTTCAAAA 389
QY
261 TCACAAAAGATCAAGGTGTGCTGTGATCTCTTAAACCAATTTGTGAAAAGATTGAAAG 320
Db
390 TACTGATCTGATGATTTGGCTGAACTATGAAACCATCATTTAGCAAAATAGAAAG 449
QY
321 GCCAAAGACGTTTGTGTTAAGTTTGTGTTAAATTTCTTCCGCTGACACACAACT 380
Db
450 GCCAAAGATGTGTGCTTCCCTGACTGATTAATTTTTCACCTGATCTGTGTAGCT 509
QY
381 CCAGAAAGAACTCAAGGATCTGTGCGCTGAGTGAAGCAGAGACTTGGCTCAAG 440
Db
510 ACAAGAAATATACAGATCTTGTGCTTGTGCTTGAACCTTAAAGAGACCTGTGAAAG 569
QY
441 CAGTTGACGTGTATGACACGAGCAGCTCTTGTATTTCAACATTTGTGCAATCTGA 500
Db
570 GGGTTTGACTGTGCTGACACGAGCGGCTTCTCAGTCCCATCTCTGCAGTGCGA 629
QY
501 GATTGGGATTTTGTGTTGAAGCTTTGACAGAGACCTTAGCAAAAATTAATATACC 560
Db
630 AATAGAGATTAAGATGAAACGCTGAGCCGAGAGACCTCAAGTGAAGAGATTTGCG 689
QY
561 TCAGCAAGACGACCTAGAGACAAATTCGTGAATTTCAACATPACCACTTGAACAA 620
Db
690 TGGCAGCAGCACTGCTTGAAGAAATCTAGAAATTTCCATCAGAACAGTGGCCAGAC 749
QY
621 ACCAGCAGATCAGATTTTCAAGCTCTAGAGATTTGCGCTGAGTGAAGATGTAAGAT 680
Db
750 ACTGTGATGCGGATTTTCAAGGTGCTCGAAATTTGCTCGAAAGTGAATGTAACGCAAT 809
QY
681 CCGGTTGACACCGGCAAGAGACAGGAGACAGAGATCAATCTGCGCTTCCCAACAC 740
Db
810 CAGATTTTCAATGTGCTTGAACAGGAGAGACCAAGATTTCAACTGCGATTTCCACAT 869
QY
741 GGAATTTCTAGTGTTCAGGGTTTCACTAAGATCAATGCTTCAACTGAGGCGCAAGTGC 800
Db
870 GGGTGTACTGTGTTTCAAGGACACCAAAATCAACCTTCACTGTGCTCAAGGTCCG 929
QY
801 GAAGCTGACCTTCAAGAGAAAGCTTTCTCATCAAGCTTCGCGCAGATGCCAATATGTC 860
Db
930 TAAACTAATCTTCAAGAGAAAGATTTCTTATCAATTCACAGAGTTTCATGAGACC 989
QY
861 GATACAGGATCTTGTGAATTCCTGATGCGAGCTGCGGATTTCTGAAAGCTCTTCTGAA 920
Db
990 TTAACAGGACACTTAAGAAATTTTGTGGGTAGTGAAGTAATGTAAAGATTTCTGAAA 1049
QY
921 AATCTGTGTGAACATCATGCTTCTTGTAGACTTTTGAAGAGCCCAACCAAGCCCA 980
Db
1050 GATTTGTGTGAGATCAACACTTTTGTAGACTTTTGAACCACTTAAGCAAAAGCAAA 1109
QY
981 GCCCGTCTCTTTTGAAGCGGGGATCATATTTGCGTTCAAGTGTGCGACTTCAGAGCAAGT 1040
Db
1110 AACCGCTCTTCTTCAAGCGGGGCTCTCTTCAAGATCAAGTGAAGAACTCAGAAACAACT 1169
QY
1041 TCTCGACTATGTTAAAGAGAGACATTAAGAGTGCAGTTTGAAGAGAGACAGCAA 1100
Db
1170 AGTAGATTTATTTCAAGAGAGTGAAGTAAGAAATTTCAATAGAAAGAGAGACAGCAA 1229
QY
1101 GATTCATTTATCTGAGAGCTTGTCTTCAAGCTTACAGACTGAAGATGAATTCGGAAGTCTGA 1160
Db
1230 GACCCACAGTCC-----GTTGAGCTCTGACGTGACAGACTTACAAA 1271
QY
1161 GAGATCTCAGAGAGACCAAGCTTACATTTGAGAAAGTGCAGAAATTCAGAGGGGCA 1220
Db
1272 ACAGAGCATCTCATTTCCGAGAGGATTTGAGACTCTGCTCCCATCTTCAAGCGATGC 1331
QY
1221 GAGCTCCCGGAGAGAAAGAACCGAAGTTTCCCGCGGAGAGCCGGGTGCGACCCGAG 1280
Db
1332 CTTTAACTGCTCTCTCCCTCACTGCTGCTCCCTGCGCTGCGAGATTAAAGACAG 1391

1281 CCTCGCGCGAGAGAAAGCCCGCGGATTAACAGAGAGCGGACGAGCCGCTCGCGCC 1340
QY
1392 CAGCAGCTCCCTCAACAGATCCCGAGGTTTCTTAAGTCAAGAGATCCAGCTGAGAGAGGG 1451
Db
1341 CACGAGAGAGAGAGAGAG--TGTTAAGATAGAGCCGACAGAGTAACTCAGCC 1397
QY
1452 CAGTGAAGCAGTGGCTGAGAGCCCGACACACATCGCGCCGCGCTCGCGCCCGC 1511
Db
1398 CCGGAGCCCAAGCAGAGCTCCCTGATCGGACAGTCCCTCACTTCCAGAGCTGTCTGTAA 1457
QY
1512 ACTCAGCTGTGTCAAGGCTTTTCAAGAGATCTTCAAGCTTCTCTCCAGCCGAA 1571
Db
1458 CTGCGAGGGGGAGTGGCCCTGCGCAAGTGAACCTGTCTCCAACTGAGCCCGACAC 1517
QY
1572 GAGCCCTCAGATCTGAGCCCTGCAATTCAGGTGCTT-----TGGCCACAGCTGA 1622
Db
1518 CAGCAGAGCTTCTCCCTTGTATCAGCCCGCTGTGAATGACAGAGCTGCCCCGACGAA 1577
QY
1623 ACAGGACTCATCCCACTCTGAGCCCTGTCTCAGTATGTCTGGCGAGCGGATGGA 1682
Db
1578 CGATGAGATGAGGGCCGAGAGAGATTCCTCACTGATTAAGGCTACTTATAGCTAA 1637
QY
1683 CTGCGAG--GAGCCGAGACAGAGCGGTGCTGACAGAGGCTTACTTATAGTCAA 1739
Db
1638 GGAAGTGTCTACACCGAGCAGACATATCTGAAGATCTGAAATTAATCACTTGTGTT 1697
QY
1740 AGAGATTTCTGCTTACAGAAAGCAATCCTCAAGGATTTAGAAATTAACGTTGTGTT 1799
Db
1698 TCAGAGCAGATGAGCAAG 1757
QY
1800 CCGCAGCGCAGTGTGAAG 1859
Db
1758 GAATTTTGAACCTTTGCAAAATTTCTAATATTTCTCAAGAAATTTGACCAAGACT 1817
QY
1860 CAACATCATCTCCATCTATGATTTCAAGAGGCTTCTGCGCAGTGTGAGAGAGGCT 1919
Db
1818 TGCCTGTGAGAGAGCGGCTCAATGCCCCAATCAGAG--ATTAACAAAGATGCGCA 1874
QY
1920 GGCACCTGTGAGAGAGGCGCTTCAAGAGCCACAAAGGAGAGATCAACAGATCGGGGA 1979
Db
1875 TGTATGCTGAAGAACTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1934
QY
1980 CATCTGTCTCAGAGAACTGTGCGCAGTTAAAGAGTTTACAGCTACTTCAAGAGATGA 2039
Db
1935 CGAGGCTTGTGAGGCTGTGAGATGGAATCAAGAGCTCCCGCGCTGAGAACTTCTG 1994
QY
2040 CAGGCTCTTAAGAGAAAGTGAAGAGGCTAACAGGCTGTAAAGATTTGAGAGAGCTGA 2099
Db
1995 CAGAGCTTTGAGCTGAGAGAGGTGTGTTACTACCGCTCAACACTTCTCTCGCGCC 2054
QY
2100 CAGAGATTTGAGCTGAGAGAGGTCTGCTACTTGTCTTCAACAGTTCCTCTGAAAGCC 2159
Db
2055 ACTGACACCGGCTCATGAGCTTAACAGAGGTCTGTGAGAGGCTGTGAAACACACCCGCG 2114
QY
2160 CATCAGCGGCTGTGCACTTACCGCTGTGTGCGCGCTTATGTGGACATTTACAGCCC 2219
Db
2115 GAGCCACGCGCACTTCAAGGACTGCGAGCGCTTTTGGCAGAGATCAAGAGATGTGTC 2174
QY
2220 CGGGACACTGACTGAGCTGACCTGCAATGAGCGCTTGAAGGCACTACAGAGGTGACAC 2279
Db
2175 ACAGCTTCAAGGTGACAGTATCAAGATGGAATTTCCAGAGCTGACGAACTCAAGAA 2234
QY
2280 CACACTACAGCAATTTCTCATCTGCTGGAAGAACTGACAGAGCTTACGAGCTGACAG 2339
Db
2235 AGATTTGATTTGCACTTGAACATCTTGTGTTCCGGAGAGGAGTTCATCCGTCTGGCAG 2294
QY
2340 GGAACCTGTGTGAGAGAACTCATTTGCTCTGCGAGAGGAGTTTATCTGAGGGCTG 2399
Db
2295 CCTCAGCAAGCTCTCGGGAGAGGGCTCAGAGAGGAGATTTCTTCTCTTCAACAGAGCT 2354
QY
2400 CTTTCAAGAGCTCAAGAGAGGGCTGACAGAGAGATTTTCTTCTCTCAGATAT 2459
Db

Qy	2355	CTGTGTAATACAGAGCCGGGGGCTGAGGCGCTCCAAATAGATTAAAGTCAACGGGCAAGCT	2414
Db	2460	GTGTGTTGACCAAGCAAGAAAGAGTTTGACGGAGACCAAGCACTTCGGAACTCCGGGGCTCTCT	2519
Qy	2415	CCCGCTCTATGGCATGACGATTTGAGAGAGCGAAAGACGATGGGGGGGTGCCCACTGCGCT	2474
Db	2520	TCCCTCTCCAAAGGATGCTGTGTGAAGAAAGTGAATACAGAGTGTCTGTCTTCCACACTGTCTT	2579
Qy	2475	GACCCCTCGGGGGCCAGCGGCAGTCCATCATCTGTGGCCGCGCAATTCTCGGTCCGAGATGGA	2534
Db	2580	CACCATCTACCGCGGCTCAGAAAAACAATGTGTGGCAGCCAGCACTCGGCTGAGAAAGA	2639
Qy	2535	GAATGTGGTTAGGACATCCAGATGCGCATTTGACCTGGCGGAGAAAGACAGACAGCCCGC	2594
Db	2640	GAATGTGATGTCTGACCTTGAACTCCGTCATCCAGACACCAAGAGTGGCGGTGTGACAGCGC	2699
Qy	2595	CCCTGAGTCTCTGGCCAGCA-----GGCCCCCTGACCAACAAGTCCCTGTATGAAGCCAC	2648
Db	2700	CCCTGTGCACTGCGAGGCGGCACTGTGTGCACTGTCTCCCCCAAGATCCCCCAAGAGTATC	2759
Qy	2649	CGCGGCTGACCCAGGAGTCAAGAGATGACCTTGAGCGCCTCGGCACTATGCTGAGCGGCA	2708
Db	2760	TCTG---GAGCAGGAGTCAAGAAATGATGATGCTCGGGGTGTGTCGCAAGCTCTCGAGGGGCA	2816
Qy	2709	GGCCCCGCAACCGCGGCAACACATGTGTGACATGTGTGCTGCGCAACCGCAACACAGCTCTC	2768
Db	2817	TGGCCAGACCGGGGCAACACACATGCACTGTGTGCTGATCCGGAACACACAGCTGTCTC	2876
Qy	2769	CATGTGTGACTTCAAGCATCGCAGTGGAGAAATCAGTGTGTGGAAACCTGCTGAGGAATT	2828
Db	2877	CAGGGCAGACCAACAGTGCAGCTGTCTGAGAACAGACTTTCAGGAAATATCTGCTTAAGAAATT	2936
Qy	2829	CAAAAACAGCAACGGGTGGCAGAAAGCTGTGGTGTGTTCAAAACTTCTGCTCTTCTT	2888
Db	2937	CAAAAACAGTCAATGGCTGGCAGAAAGCTCTGGGTGTCTTTACAACTTCTGTGTTCTT	2996
Qy	2889	CTACAAATCACACACAGACAAATCATCTCCCTTTCAGAGCTGCTCTTCTCGGCTATCGCT	2948
Db	2997	CTACAAAATCTCATAGATGATGACTACCCACTGGCCAGCTCCCGCTGGCGGCTTACAGCGT	3056
Qy	2949	CACCATCCCTCTAGTCCGAGAACATCCAAAGAACTACGTGTTCAAAGCTGCACCTTCAA	3008
Db	3057	GAGCATCCCAAGGAGGCGCATGCGATACACAAAGACTATGTTTTCAAGCTCCAGTTCAA	3116
Qy	3009	GTCCCAAGCTCTACTACTTCAAGGAGCGGAAAGCGATACAGCTTCGAAAGGTGATGAGAT	3068
Db	3117	ATCCCAAGCTCTACTCTTTCGCGGCTGAGAGCAATACATTTGAAAGGTGATGAGAT	3176
Qy	3069	GATCCGAGTGCACACAGCTCTGCTTCGCGAGCCCT	3103
Db	3177	GATCCAGGGGGCCAGAGCTCAGCCCGGAGGGGCC	3211
RESULT 6			
US-10-342-887-1536			
; Sequence 1536, Application US/10342887			
; Publication No. US20040058340A1			
GENERAL INFORMATION:			
APPLICANT: Dai, Hongyue			
APPLICANT: He, Yudong			
APPLICANT: Linsley, Peter S.			
APPLICANT: Mao, Mao			
APPLICANT: Roberts, Christopher J.			
APPLICANT: Van 't Veer, Laura Johanna			
APPLICANT: Van de Vijver, Marc J.			
APPLICANT: Bernards, Rene			
TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients			
FILE REFERENCE: 9301-188-999			
CURRENT FILING DATE: US/10/342, 887			
PRIOR APPLICATION NUMBER: 2003-01-15			
PRIOR APPLICATION NUMBER: 60/298, 918			
PRIOR FILING DATE: 2001-06-18			
PRIOR APPLICATION NUMBER: 60/380, 710			

	PRIOR FILING DATE: 2002-05-14	
	PRIOR APPLICATION NUMBER: 10/172,118	
	PRIOR FILING DATE: 2002-06-14	
	NUMBER OF SEQ ID NOS: 2699	
	SEQ ID NO 1536	
	LENGTH: 3997	
	TYPE: DNA	
	ORGANISM: Homo sapiens	
	US-10-342-887-1536	
Query Match	33.1%; Score 1037.4; DB 8; Length 3997;	
Best Local Similarity	60.2%; Pred. No. 6,7e-302;	
Matches 1864; Conservative	0; Mismatches 1186; Indels 45; Gaps 7,	
Oy	21 GCCGACCCAGAGATCAGCATGGGGGCCCGGAAATTCCGGGAGTCACTACTTTGAACG 80	
Dd	150 GCAGCTGCAGAGATGCGCTGGGTGCCAACCCTGTGGGAGTTAGCACCTCTTGAGCC 209	
Oy	81 TGGACAGAAGCCCCCCCACAACCTTCAGGAAAACTCGTGTTCATCAAATAATCCAGATGCT 140	
Dd	210 TGGCGAGACTCTCTTCCAGAAATGCAAGAAAGAACCTGCACCTCAGAGTAAGCTCT 269	
Oy	141 GGATGACACCCAGAGGAGGATTTGAAGTTCCAACAAGAGCTCTCGGGAAGTGCTCTGSA 200	
Dd	270 GGAACAACCCATGGAAATATTGACATTGAGCTTAATGCGATGCGCCAGGATTAATGAC 329	
Oy	201 TGCAATTGCAACCACTCAACCTCGTGAAGTGATCTATTTTGGCCTCGAATTCTGA 260	
Dd	330 ACAAGTGTGAAGGGTTTTAACTGTGATGATGATCTACTTTGGGAGTAGAGTTTCAAAA 389	
Oy	261 TCACAAAAAGATCAACGGTGTGGCTGATCTCTTAAACCCATTGTGAACGATTAGAG 320	
Dd	390 TACTCAGTCTTACGTGAATTTGGCTTGAACCTATGAACCCATCATTTAGGCAAAATACGAAG 449	
Oy	321 GCCAAGACGCTGTGTGTAGTTTGTGTGAATTTCTTCCGCTGACCAACACAACT 380	
Dd	450 GCCAAAGATGTGGTGCTCGCTGACCTGTAATAATTTTCCACTGTATCTTGCTGACT 509	
Oy	381 CCAAGAAAGATCAACAAGTACTCTGTTCCGCTCAGGTGAAGAGCACTTGGCTCAAG 440	
Dd	510 ACAGAAAGAAATATACAGATATCTGTTGGCTTGCACCTTAAGAGAGACTGTGGAAGA 569	
Oy	441 CAGGTGACGTGTAATGACACACAGCGCAGCTCTCTGATTTCAACATTTGTGCAATCTGA 500	
Dd	570 GCGTTTGACCTGTGCTGACACCAAGCGGCCCTTCTCAGCTCCATCTCTGCACTGCGA 629	
Oy	501 GATTGGGGATTTTATGATGAAGCTTGGACAGAGACACTTACCAAAAAATAATACATCC 560	
Dd	630 AATGAGGATTAACATGATGAAGAGCTGAGACCGAGACCACTCAAAATGAACGATATTGGCC 689	
Oy	561 TCACAGAAAGCACTAAGAGCAAAATCGTGAATTTTCAACATTAACCACTTGAACAAC 620	
Dd	690 TGGCCAGACGACCTGCTTGAAGAAATCTAGAAATTCATATAGAGACCTGTGGGCCAGAC 749	
Oy	621 ACCAGCAGAAATCAGATTTCCAGCTCTAGAGATTTGCCGCTGGCTAGATGATGAAAT 680	
Dd	750 ACTGTGAGATGCGGATTTCCAGGTGCTCGAAATTTGCTCGAAAGTTGGAAATGTACGGCAT 809	
Oy	681 CCGGTGCAACCCGCCCAAGAGCAGGGAAGCGACGAAGATCAATCTGGCCGTTGCCAACAC 740	
Dd	810 CAGATTTTCACTGGCTTCTGAACAGGAAAGCAACAAATTCATACCTGGAGTTTCCACAT 869	
Oy	741 GGGAAATCTAGTGTTCAGGGTTTCACTAAGATCAATGCTTCAACTGGGCCAAGGTGCG 800	
Dd	870 GGGGTACTCTGTGTTCCAGGGACACCAAAATTAACATTTCAACTGTGTTCCAAAGTCCG 929	
Oy	801 GAAGCTGAGCTTCAAGAGAAAGCGCTTCTCATCAAGCTCCGGCCAGATGCAATAGTGC 860	
Dd	930 TAAACTAAGCTTCAAGAGAAAGATTTCTTATCAAACTTCATCACAAGGTTTCAATGACC 989	
Oy	861 GTACCAAGATCTCTTGAATTCCTGATGCGCAGTCCGGATTTCTGCAAGTCTTCTGGA 920	
Dd	990 TTACACAGACACATTAAGATTTTGTGGGTAGTAGATGAATGATGAAGAACTTCTGGA 1049	

QY 921 AATGTGTGAACATCATGCGCTTTTGAAGCTTTTGAAGGCCAAACAAAGCCCA 980
DB 1050 GATTGTGTGTGATACACCTTTTGAAGCTTTTGAAGCTTTTGAAGGCCAAACAA 1109
QY 981 GCCCGTCTCTTGAAGCGGGGTGATCATTTTGGTTCAAGTGTGCGATCAGAGAGGT 1040
DB 1110 AGCCGTCTTTCAGCGGGGTCTCTCTTCAAGTGTGAGAGAACTCAGAAACAAT 1169
QY 1041 TCTGACTTGTAAAGAAAGAGACATTAAGAGTGTGAGTTGAAAGAGACAGCA 1100
DB 1170 AGTAGTATTTTCAAGAGAGTGTGAGAGAAATTCATTAAGAAAGAGACAGCA 1229
QY 1101 GATTCTTATCCGAGCGCTTGCTTCAAGCTTACAGAACTGAATTCGAAAGTGTGGA 1160
DB 1230 GACCCACAGTCC-----GTTCAAGCTTGAAGTGTGAGAGCTTACCAAA 1271
QY 1161 GGAGTCTCAGAGAGACAGCCTTACATTTGAGAGAGTGTGAGAGTGTGAGAGAGCA 1220
DB 1272 ACAGAGCATCTCATTCCTCGAGAGGATTTGAGAGTCTGCTGCTCCCATCTTCAAGCAATGC 1331
QY 1221 GAGCTGCGGAGAGAAAGAACCGAAGTTTCCGCGGAGAGCGGGGTGCAAGCCGAG 1280
DB 1332 CTTTTAACTGCTCTCTCCCTCACTGTGCTCCCTGCGCTGCGCAGAGTTTAAAGACAG 1391
QY 1281 CCTGCGCGAGAGAGAGCGCGCGGATTAAGAGAGCGAGAGCGCGCTCGCGGC 1340
DB 1392 CAGAGCTCTCTCAGATTCCTCAAGTTCTTACGTCAAGAGTGTGAGAGTGTGAGAGAGG 1451
QY 1341 CACGAGAGAGAGAGAGAGG---TGTGTAAGATGAGAGCCAGAGAGTAACTCAGCC 1397
DB 1452 CAGTGAAGAGTGTGTGAGAGCGCCGACACCATGCGCCAGCGCTCGCGGCCCGCCG 1511
QY 1398 CCGGAGCCAGAGCAGAGCTCTCTGATGTGAGAGTCTTCACTTCCGAGCTGTGTGA 1457
DB 1512 ACTCAGAGCTGTGTGAGAGCTTTTCAAGAGTGTGTGAGAGCTTCTCTCCAGCGGAA 1571
QY 1458 CTGCGAGGGGAGTGTGAGCGCTGCAAGTGTGAGCTTGTCTCCCACTGAGAGCGCCGAC 1517
DB 1572 GAGCGCTGTGATGTGAGAGCTTCAAGTGTGCT-----TGGGCCAGCTGA 1622
QY 1518 CAGCAGAGCTCTCTCTTATGAGCGCGCTGTGATGAGCAGAGCTGTGCGCGAGCA 1577
DB 1623 ACAGGGCTATCCCACTCTGTGAGCGCTGTCTCAGTGTGTGAGAGCGCGAGTGA 1682
QY 1578 CGATGAGATGAGAGCGCGAGAGAGATTTCCCACTGATTAAGCTTACTTACTTAA 1637
DB 1683 CTGCGAG---GAGCCAGACACAGAGCGGTGTGAGAGAGCTTACTTCAATGTCAA 1739
QY 1638 GGAAAGTGTACACCGAGAGAGCATATCTGAGAGATCTGAGATTAATCACTTGTGTG 1697
DB 1740 AAGAGATTCGTGACAGAGAGAGACATCTCAGAGATTTAGAGATTTACCTGTGTG 1799
QY 1698 TCAGAGCAGATGAGAGAGAGAGAGCGCATGTGCGAGAGCACTGAAAGTCTCATATCCC 1757
DB 1800 CGCAGCGAGTGTGAGAGAGAGAGCGCATGTGCGAGCTGTGATGACCTGTCTTCTC 1859
QY 1758 GAATTTTGAACCTTTGACAAATTTTCACTAATTTTTCAGAGAAATTTGAGAGAGAT 1817
DB 1860 CAACATCGATCCCATATGATGATTCACAGAGGCTTCTGCGGAGTGTGAGAGAGGCT 1919
QY 1818 TGCCCTGTGAGAGAGCGCTCAATATGCCAAATCAGAG---ATTACAAAGATCGGCA 1874
DB 1920 GGCATCTGTGAGAGAGCGCTCTCAAGCCACACAAAGAGAGTCAACAGATTCGGGA 1979
QY 1875 TGTGATGTGAAGACATTCAGGAGTGAAGACCTGTGCGGCTCACCTGTGAGAGCAG 1934
DB 1980 CATCTGTCTCAGAGACATGTGCGCAGTTTAAAGAGTTTACAGTACTTCCAAAGACATGA 2039
QY 1935 CGAGGCTTGTGAGAGCGCTGTGAGAGATGAGATCAAGAGCTCCCGCGCTGTGAGAGACTTCTG 1994
DB 2040 CGAGGCTTAAAGAGAGAGAGGCTACCAAGCCTGTAAAGATGTGAGAGAGTGA 2099

QY 1995 CAGAGACTTTGAGCTGAGAGAGTGTGTTACTTACCGCTCAACACCTTCTCTGCGGC 2054
DB 2100 CAGAGAGTTTGTGCTGAGAGAGTGTGCTACTTGTGCTTCAACAGTGTCTGTGAAGGC 2159
QY 2055 ACTGACCGGCTCATGACCTACAGAGAGTCTGTGAGCGGCTGTGCAACACCGCGC 2114
DB 2160 CATTCAGCGGCTGTGACATCCGCTGTGTGCGCGCTGTGTGAGAGATTAACAGGCC 2219
QY 2115 GAGCAGCGGCTTCAAGGATCTGCGAGCGCTTTTGGCAGAGATCATCGAGAGTGTGCG 2174
DB 2220 CGGAGCAGATGACTAGCTGATGCTGATGAGCGCTGTGAAGCTCAACAGAGTGAACAC 2279
QY 2175 ACAGCTTCAAGGATGATGATGAGAGATTTTCCAGAGCTGACGAATCAAGAA 2234
DB 2280 CACATGACAGCATTTCTCATCGGTGTGAGAACTGTGAGAGTCAACGAGAGCTGACGCG 2339
QY 2235 AATTTTGAATTTGATGACATTTCTGTGTTCGGGAGAGAGTTCATCTCTGTGCGAG 2294
DB 2340 GAGCTGTGTGAG 2399
QY 2295 CCTCAGAGAGCTCTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2354
DB 2400 CTTTCAAGAGCTCAAG 2459
QY 2355 CCTGCTATACAGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2414
DB 2460 GTTGTGTGTACAG 2519
QY 2415 CCGGCTTATGAGATGAGAGATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2474
DB 2520 TCCCTTCCAG 2579
QY 2475 GAGCTTCCGAG 2534
DB 2580 CACCATTTAG 2639
QY 2535 GAAAGTGTGAGAGATTCAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2594
DB 2640 GAAAGTGTGAG 2699
QY 2595 CCTGAGTCTGTGAG 2648
DB 2700 CCTGAGCTGAG 2759
QY 2649 CGGAGCTGAG 2708
DB 2760 TCTG---GAGCAG 2816
QY 2709 GAGCGCGAG 2768
DB 2817 TGGCAG 2876
QY 2765 CATGTGAGATTCAGATGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2828
DB 2877 CAGGAG 2936
QY 2829 CAAAAAG 2888
DB 2937 CAAAAAG 2996
QY 2889 CTACAAATCAG 2948
DB 2997 CTACAAATCAG 3056
QY 2949 CACCATCCCTTGTGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3008
DB 3057 GAGCATTCCTCAG 3116
QY 3009 GTCCAGAGTCTACTTCTTCAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3068
DB 3117 ATCCAGAGTCTACTTCTTCAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3176
QY 3069 GATCCGAG 3103

Db 3177 GATCCAGGGGGCCAGCAGCTCAGCGCGGAGGGCCC 3211

RESULT 7
US-10-802-432-16
; Sequence 16, Application US/10802432
; Publication No. US20040185489A1
; GENERAL INFORMATION:
; APPLICANT: Hoffmann-La Roche Inc.
; TITLE OF INVENTION: Transcriptional Activity Assay
; FILE REFERENCE: 21574
; CURRENT APPLICATION NUMBER: US/10/802,432
; CURRENT FILING DATE: 2004-03-17
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 16
; LENGTH: 3997
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: KIA00793
; LOCATION: (1)..(3997)
; OTHER INFORMATION: GenBank accession No. AB018336
US-10-802-432-16

Query Match 33.1%; Score 1037.4; DB 9; Length 3997;
Best Local Similarity 60.2%; Pred. No. 6,7e-302;
Matches 1864; Conservative 0; Mismatches 1186; Indels 45; Gaps 7;

Qy 21 GCCGACCCCGAGATCACTGAGGCGGCCGGAATAATTCGGGGATCAGTACTTGGAAAG 80
Db 150 GCAGACTGCGAGGATGCGCTTGGTCCCAAGACCCCTGTGGAGTTAGCACCCTTGAGCC 209
Qy 81 TGGACAGAAAGCCGCCCAACACCTTCAGAAAACTGTGTCCATCAATCAATTCAGATGT 140
Db 210 TGGGACAGACTCTCTTGGCCAGAAATGCAAGAAAGACCTGACCTCAGAGTAAAGCTGCT 269
Qy 141 GGAATGACCCCGAGAGGAGCTTTGAAATTCACAAAGAGCTCTGGGAGGTGCTCTGGA 200
Db 270 GGCACACACCATGGAATATTTGACATTGAGCCTAATGCGATGCGCAGGTATTTACTGAC 329
Qy 201 TGCAGTTTGCACACACTCAACCTCGTGAAGGTAATTTTGGCTCGAGTTTCTCTGA 260
Db 330 ACAAGTGTGAAAGCTTTAAACCTGTGAATGACTTCTGGGATGAGATTTCAAAA 389
Qy 261 TCACAAAAGATCAGCGTGTGCTGATCTCTAAAACCCATTGTGTAACAGATTAGAA 320
Db 390 TACTCAGTCTTCTGATTTGGCTTGAACCTATGAAACCCATCATTTAGCAAAATAGAA 449
Qy 321 GCCAAAGACGTTGTTGTTAAAGTTTGTGTAATTTCTTCCGCTGACACACACACT 380
Db 450 GCCAAAGATGTGTGCTTGCCTAGCTGTAATAATTTTTCACCTGATCTGTGCTAGCT 509
Qy 381 CCAAGAAGACTCAGAGGTACCTGTTGCGCTGAGGAGTGAAGAGGACTTGGCTCAAG 440
Db 510 ACAAGAAATTAACAAGATATTTGTTGCTTGCACCTTAAGAGAACCTGTGGAAGA 569
Qy 441 CAGGTGACGTGTATGACACAGCGAGCTCTCTGATTTCAACATTTGCAATTTGA 500
Db 570 GCGTTTGACCTGTGCTGACACACAGCGCCCTTCTCAGCTCCCATCTCTGCACTGGA 629
Qy 501 GATTGGGATTTTGTATGAAAGCTTTGACAGAGACCTTAGCAAAAAATTAATACATACC 560
Db 630 AATAGAGATTAAGATGAAACGCTGACGAGAGCACTCAAAAGTGAACAGATATTTGGCC 689
Qy 561 TCAGCAAGACGACTAGAGCAAAATCTGAAATTTCAACATTAACCAATTGGAACAAC 620
Db 690 TGGCCAGCAGCATCTGCTTGAAGAAAGATCTAGAAATTTCAACAGAGACCTGGGCCAGAC 749
Qy 621 ACCAGCAGAAATCAATTTCCAGCTCTTGAAGATTTCCGCTCGCTAGAGATGATGAAAT 680
Db 750 ACCTGCTGAGTCCGATTTCCAGGCTGCTGAAATTTGCTGAAAGTTGAAATGTACGGCAT 809

Qy 681 CCGTTGCAACCGGCCAAGGACAGGAAAGGCAAGAAATCATCTGGCGTGGCAACAC 740
Db 810 CAGATTTCAATGAGCTTCTGACAGGAGAGAACCAAGATTTCACTGGCACTTCCCAT 869
Qy 741 GGGAAATTTAGTGTTCAGGGTTTCACTAAGATCAATGCTTCACTGGCCAAAGTGC 800
Db 870 GGGGTGATCTGTGTTCAGAGGACACACAAATTCACACTTTCACTGTGTCAGAGGTCC 929
Qy 801 GAAGCTGAGCTTCAAGAGGAGCGCTTTCATCAAGCTCCGGCAGATGCCAATAGTGC 860
Db 930 TAAACTAGCTTCAAGGAGGAAAGATTTCTATCAAACTTCACAGAGGTTCAATGACC 989
Qy 861 GTACCAAGATACCTTGAATTCGATGCGCAGTCCGGAATTTGCAAGGCTTTGGA 920
Db 990 TTACAGAGACATTAAGAAATTTTGTGGGTAGAGATGAATGTAGAACTTCTGGA 1049
Qy 921 AATCTGTGTAACATCATGCTCTTCTTACCTTTTGAAGAGCCCAACCAAGCCCA 980
Db 1050 GATTGTGTGATATCAACCTTTTGTAGACTTTTGAACCAACTTAGCCAAAGCAAA 1109
Qy 981 GCCGCTCTTTTAAAGCGGGGCTCATCTTGGTTCACTGTGCTGAGTCAAGACAGGT 1040
Db 1110 AGCCGCTTCTTCAAGCGGGGCTCTCTTCAATACAGTGAAGAACTCAGAAACAAT 1169
Qy 1041 TCTGACTATGTTAAAGAAAGAGACATAAAGGTGCTGTTGAAGAGGACAGCA 1100
Db 1170 AGTATGATTTATTAAGACAGTGAATGAAGAAATTCATATGAAGAAAGGACAGCA 1229
Qy 1101 GATTCATTTATCTCGAGAGCTTCTTCAAGCTTACAGCTTACAGAAATTCGAAATGCTGA 1160
Db 1230 GACCCACAGTCC-----GTTGAGCTCTGACAGACCTTACCTAAC 1271
Qy 1161 GCAGTCTCAGCAGAGACACAGCTTACATTTGAAGAAAGTCCGAATCTCAGAGGGCCA 1220
Db 1272 ACAGACATCTCATTTCCCGAGGAGATTTGAAGACTCTGCTCCCACTTTCACAGATGC 1331
Qy 1221 GAGCTCCGCGAGGAAAGAACCAAGGTTTCCGCGGGAGCCGGGGTCCGACCCGAG 1280
Db 1332 CTTTTACTGTCTCTCTCCCTCACTGTGTCCCTCTGCGCTCCAGAGTTTAAGACAG 1391
Qy 1281 CCTGCGCGAGGAGAAAGCCCGGGTTAAACAAGAGCGAGAGCCGCTCGCGGCC 1340
Db 1392 CAGCAGCTCCTTCAAGATCCCAAGTTTCTTACGTCAAGATGTCAGCTGCAAGAGGCG 1451
Qy 1341 CACGAGAAAGAGAGAGAG--TCTTAAAGATAGAGCCAGCAAGATTAACCTACGCC 1397
Db 1452 CAGTGAAGCAGTGGCTGAGAGCCCGACACACATCGGCCCAAGCCCTCGGGCCCGCG 1511
Qy 1398 CCGCAGCCCAAGACAGGCTCCCTGACTGCACTGCACTTCCAGAGCTGTCTGTGA 1457
Db 1512 ACTCCAGCTGTGTCCAGGCTTTTCCAGAAAGATCTCAGCTTCTCTCCACCGGAA 1571
Qy 1458 CTCGAGAGGGAGTGGCCCTGCAACGTAACCTTGTCTCCCACTTGAGCCCGACAC 1517
Db 1572 GAGCCCTTGAATCTAGAGCTTCAATTTCAAGTGCCTT-----TGGGCCAGCTGA 1622
Qy 1518 CAGCAGGCTCTTCCCTTATGAGCCGCTGTGTAATGACAGGCTGCCCCGAGAGGA 1577
Db 1623 ACAGGCTCATATCCCACTCTGAGCCCTGCTCATGTATGCTGAGGAGAGCCGGATGGA 1682
Qy 1578 GATGAGAGATGAGAGGCGGAGAGAAATTCCTCAACTGATTAAGCGATCTTCAAGCTTA 1637
Db 1683 CTGCGAG--GAGCCCAAGACAAAGCGGTGCTGAGAGAGGCTTCACTTCAATGCA 1739
Qy 1638 GGAAGTGTCTACCAAGAGGAGCAATATCTGAAGATCTGAAATTTATCATCTTGTGTT 1697
Db 1740 AGAGATTTCTGCTACAGAAAGAAATACCTCAAGATTTTAAGATTTATACGTTGGTT 1799
Qy 1698 TCAGAGCAAGTGAAGAAAGAGAGCGCATGCGGAAAGCACTGAAATGTCATATTTCC 1757
Db 1800 CCGCAGCGCAGTGTGAAGAGAGACCATGCTGGAATCTGATGACGCTGCTCTTCTC 1859

Db 433 CCCGATCCCCCAAGATATCTCTG---GAGCAGGAGTCAGAAATGATGTCGGGGT 489
Qy 2686 TCCGCGCATTCGCTGGAGCCCGCAGCCCGCAGCAACAAATGGTCAGTTCG 2745
Db 490 GTCGCGAGCTCCCTGGAGGGGCGATGCGACACCGGGCCAAACAAATGCGACGTG 549
Qy 2746 TGGCACCAGCAACACAGCGTCTCCATGATGGAATTCAGCATTCGAGTGGAAATCA 2805
Db 550 TGGTACCGGAACACAGCGTCTCCAGGGCAGACACAGTGCAGTGTGAGAACCACTT 609
Qy 2806 TCTGGAACCTCTGTGAGAAATTCAAAACAGCAACGGGTGGCAGAAAGTGGGTG 2865
Db 610 TCAGGATATCTGTGTAAGAAAGTTCAAAAACAGTCATGAGTGGCAGAAAGTGGT 669
Qy 2866 TTCACAAACTTGTGCTGCTGTTCTTCTACAAATCAACAGGACATATCCCTTGCAC 2925
Db 670 TTACCAACTTGTGTTGTTCTTCTACAAACCTATAGATGCTACCCACATGGCCAG 729
Qy 2926 CTGCTCTGCTCGGCTACCTGCGCTCAGCATCCCTCTGAGTCGAGAACATCCAGAA 2985
Db 730 CTCCTGCTGCTGGCTACAGCTAGCATCCCGAGGAGCCGATGCGATACAAAGAC 789
Qy 2986 TACGTTTCAAGCTGCACTTCAAGTCCACGCTTACTTACTTACAGGGCGAAAGCAG 3045
Db 790 TATGTTTCAAGCTCCAGTTCAATCCCAAGTCTTACTTCTTCCGGGCTGAGAGCA 849
Qy 3046 ACGTTCAAAAGGTGATGAGAAAGTATCCGAGTCCACAGCTTGTCTTGCACCCC 3103
Db 850 ACAATTTAAAGGTGATGAGAGTATCCAGGGGCGCAGCTCAGCCGGAGGGGCC 907

RESULT 9

US-09-918-995-8794
; Sequence 8794, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: FROM VARIOUS CDNA LIBRARIES
; CURRENT APPLICATION NUMBER: US/09/918, 995
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235, 076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 8794
; LENGTH: 393
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(393)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-8794

Query Match 12.3%; Score 385.6; DB 3; Length 393;

Best Local Similarity 99.0%; Pred. No. 3,4e-105;
Matches 388; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 862 TACCAGATACCTTGAATTCCTGATGCGAGTGGGATTTCTGCAAGTCCCTTGGAAA 921
Db 1 TACCAGATACCTTGAATTCCTGATGCGAGTGGGATTTCTGCAAGTCCCTTGGAAA 60
Qy 922 ATCTGTGTGACATCATGCTCTTTAGACTTTTGAAGAGCCCAACCAAGCCCAAG 981
Db 61 ATCTGTGTGACATCATGCTCTTTAGACTTTTGAAGAGCCCAACCAAGCCCAAG 120
Qy 982 CCGCTCTCTTTAGCCGGGGGTCAATCTTGGTTCAGTGTGGAATCAGACAGGTT 1041
Db 121 CCGCTCTCTTTAGCCGGGGGTCAATCTTGGTTCAGTGTGGAATCAGACAGGTT 180
Qy 1042 CTCGACTATGTAAAGAAAGGACATTAAGAAAGTGCAGTTGAAAGGACACAGCAG 1101

Db 181 CTCGACTATGTAAAGAAAGGACATTAAGAAAGTGCAGTTGAAAGGACACAGCAG 240
Qy 1102 ATCTGTGTGACATCATGCTCTTTAGACTTTTGAAGAGCCCAACCAAGCCCAAG 1161
Db 241 ATCTGTGTGACATCATGCTCTTTAGACTTTTGAAGAGCCCAACCAAGCCCAAG 300
Qy 1162 CAGTCTCAGCAGACACAGCTTACATTTGAGAAAGTCCGAAATCTCCAGGGGCGCAG 1221
Db 301 CAGTCTCAGCAGACACAGCTTACATTTGAGAAAGTCCGAAATCTCCAGGGGCGCAG 360
Qy 1222 AGCTGCGGCGAGAAAGGAAAGGAGTTTC 1253
Db 361 AGCTGCGGCGAGAAAGGAAAGGAGTTTC 392

RESULT 10

US-09-967-768A-104/c
; Sequence 104, Application US/09967768A
; Patent No. US20020150877A1
; GENERAL INFORMATION:
; APPLICANT: Augustus, Meena
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signature
; FILE REFERENCE: 689290-72
; CURRENT APPLICATION NUMBER: US/09/967, 768A
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US/60/236, 109
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236, 034
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236, 111
; PRIOR FILING DATE: 2000-09-28
; NUMBER OF SEQ ID NOS: 325
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 104
; LENGTH: 506
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-967-768A-104

Query Match 11.4%; Score 358.4; DB 3; Length 506;

Best Local Similarity 99.7%; Pred. No. 6,7e-97;
Matches 359; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2776 GACTTCAGATCCGAGTGAAGATCACTTCTGGAACCTGCTGAGAAATTCAAAAC 2835
Db 506 GACTTCAGATCCGAGTGAAGATCACTTCTGGAACCTGCTGAGAAATTCAAAAC 447
Qy 2836 AGCAACGGGTGGAGAAAGTGTGGTGTGCAAACTTCTGCTCTTCTACAA 2895
Db 446 AGCAACGGGTGGAGAAAGTGTGGTGTGCAAACTTCTGCTCTTCTACAA 387
Qy 2896 TCAACACGAGACATCATCCCTTGGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2955
Db 386 TCAACACGAGACATCATCCCTTGGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 327
Qy 2956 CCGTCTGAGTCCGAGAAATCCAGAAAGTCAAGTGTCAAGTGTCAAGTGTCCAC 3015
Db 326 CCGTCTGAGTCCGAGAAATCCAGAAAGTCAAGTGTCAAGTGTCCAC 267
Qy 3016 GTCTACTACTTCAAGGGGGAAGCGAGTACAGCTTGAAGGAGTGAAGTGAAGTCCG 3075
Db 266 GTCTACTACTTCAAGGGGGAAGCGAGTACAGCTTGAAGGAGTGAAGTGAAGTCCG 207
Qy 3076 AGTGCACACAGCTCTGCTGCGACCCCAAGTGTGAGCCCAAGAGTCTTGTGTAT 3135
Db 206 AGTGCACACAGCTCTGCTGCGACCCCAAGTGTGAGCCCAAGAGTCTTGTGTAT 147

RESULT 11

US-10-843-6249/c
; Sequence 6249, Application US/10843641A

```
/ Publication No. US20050064454A1
/ GENERAL INFORMATION:
/ APPLICANT: Avalon Pharmaceuticals, Inc.
/ TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
/ FILE OF INVENTION: Signature Gene Setc
/ FILE REFERENCE: 689290-189
/ CURRENT APPLICATION NUMBER: US/10/843,641A
/ CURRENT FILING DATE: 2004-05-12
/ PRIOR APPLICATION NUMBER: US/09/873,367
/ PRIOR FILING DATE: 2001-06-05
/ PRIOR APPLICATION NUMBER: US/09/954,531
/ PRIOR FILING DATE: 2001-09-18
/ PRIOR APPLICATION NUMBER: US/09/954,456
/ PRIOR FILING DATE: 2001-09-25
/ PRIOR APPLICATION NUMBER: US/09/962,436
/ PRIOR FILING DATE: 2001-09-25
/ PRIOR APPLICATION NUMBER: US/09/962,832
/ PRIOR FILING DATE: 2001-09-25
/ PRIOR APPLICATION NUMBER: US/09/964,824
/ PRIOR FILING DATE: 2001-09-27
/ PRIOR APPLICATION NUMBER: US/09/967,768
/ PRIOR FILING DATE: 2001-09-28
/ PRIOR APPLICATION NUMBER: US/09/968,007
/ PRIOR FILING DATE: 2001-10-02
/ PRIOR APPLICATION NUMBER: US/09/969,347
/ PRIOR FILING DATE: 2001-10-02
/ PRIOR APPLICATION NUMBER: US/09/969,708
/ PRIOR FILING DATE: 2001-10-03
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 8447
/ SOFTWARE: PatentIn version 3.0
/ SEQ ID NO 6249
/ LENGTH: 506
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ US-10-843-641A-6249

Query Match
Best Local Similarity 99.7%; Score 358.4; DB 10; Length 506;
Matches 359; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2776 GACTTCAGCATCGCAGTGGAGAAATCAAGTGTCTGGAACCTGCTGAGAAATTCAAAAC 2835
DB 506 GACTTCAGCATCGCAGTGGAGAAATCAAGTGTCTGGAACCTGCTGAGAAATTCAAAAC 447
QY 2836 AGCAAGGGTGGAGAAAGCTGGTGGTGTCTTCAAACTTCCCTGTTCTTCAAAA 2895
DB 446 AGCAAGGGTGGAGAAAGCTGGTGGTGTCTTCAAACTTCCCTGTTCTTCAAAA 387
QY 2896 TCACACACAGGACATATCCCTTGGCAGCTGCTCTGCTCGGCTACTGCTCAACATC 2955
DB 386 TCACACACAGGACATATCCCTTGGCAGCTGCTCTGCTCGGCTACTGCTCAACATC 327
QY 2956 CCCTCTGAGTCCGAGAAATCATCGAAAGACTACGTGTTCAGAGTTCGACTCAAGTCCAC 3015
DB 326 CCCTCTGAGTCCGAGAAATCATCGAAAGACTACGTGTTCAGAGTTCGACTCAAGTCCAC 267
QY 3016 GCTTACTACTTCGAGGCGGAAAGCGAGTCAAGTGTGAAAGGAGTGAAGTGAATCCCG 3075
DB 266 GCTTACTACTTCGAGGCGGAAAGCGAGTCAAGTGTGAAAGGAGTGAAGTGAATCCCG 207
QY 3076 AGTGCCACAGCTCTGCTCGCGACACCCACAGTGTGAGCCACAAAGAGTCTTTGTGTAT 3135
DB 206 AGTGCCACAGCTCTGCTCGCGACACCCACAGTGTGAGTCAAAAGTCTCTTGTGTAT 147

RESULT 12
US-09-960-352-154
/ Sequence 154, Application US/09960352
/ Patent No. US20020137139A1
/ GENERAL INFORMATION:
/ APPLICANT: Warren, Wesley C.
/ APPLICANT: Tao, Nengbing
```

```
/ APPLICANT: Byatt, John C.
/ APPLICANT: Mathalagan, Nagappan
/ TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
/ FILE OF INVENTION: MUSCLE AND FAT DEPOSITION
/ FILE REFERENCE: 16511.006/37-21 (10298) C
/ CURRENT APPLICATION NUMBER: US/09/960,352
/ CURRENT FILING DATE: 2001-09-24
/ NUMBER OF SEQ ID NOS: 15112
/ SEQ ID NO 154
/ LENGTH: 379
/ TYPE: DNA
/ ORGANISM: Bos taurus
/ OTHER INFORMATION: Clone ID: 01-LIB34-020-Q1-E1-A9
/ US-09-960-352-154

Query Match
Best Local Similarity 9.5%; Score 297; DB 3; Length 379;
Matches 331; Conservative 0; Mismatches 35; Indels 3; Gaps 1;

QY 1587 TGAGGGCCGAGAGAGATTCCTCAACTGATTAAGGCTACTTCATAGCTTAAGAGTGTG 1646
DB 11 TGGGGCCCGAGAGAGATTCCTCAACTGATTAAGGCTACTTCATAGCTTAAGAGTGTG 70
QY 1647 TACCAACGAGCGAATATCTGAAGATCTGAAGTATCACTTGTGCTTTCAGAGCAC 1706
DB 71 CACCACTGAGAGACATATCTGAAGATCTGAAGTATCACTTGTGCTTTCAGAGCAC 130
QY 1707 AGTGAGCAAGAGAGAGAGATTCCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1766
DB 131 AGTGAGCAAGAGAGAGAGATTCCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 190
QY 1767 ACCCTTGCAAAATTTCTAATTTCTCAAGAAATTTGAGCAAGAGAGAGAGAGAGAGAG 1826
DB 191 ACCCTTGCAAAATTTCTAATTTCTCAAGAAATTTGAGCAAGAGAGAGAGAGAGAGAG 250
QY 1827 GGAAGCCGCTCAATGCCCAATCA---GAGATTACCAAGATTCGAGTCTATGCT 1883
DB 251 GGAAGCCGCTCAATGCCCAATCAAGAGATTTACCAAGATTCGAGTCTATGCT 310
QY 1884 GAAAGACATTCAGGGCAGTGAACACCTGCGGCTCAACCTGTTGAAAGCAGAGAGGCTT 1943
DB 311 GAAAGACATTCAGGGCAGTGAACACCTGCGGCTCAACCTGTTGAAAGCAGAGAGGCT 370
QY 1944 GAGAGGCTT 1952
DB 371 GAGAGGCTT 379

RESULT 13
US-09-960-352-156
/ Sequence 156, Application US/09960352
/ Patent No. US20020137139A1
/ GENERAL INFORMATION:
/ APPLICANT: Warren, Wesley C.
/ APPLICANT: Tao, Nengbing
/ APPLICANT: Byatt, John C.
/ APPLICANT: Mathalagan, Nagappan
/ TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
/ FILE OF INVENTION: MUSCLE AND FAT DEPOSITION
/ FILE REFERENCE: 16511.006/37-21 (10298) C
/ CURRENT APPLICATION NUMBER: US/09/960,352
/ CURRENT FILING DATE: 2001-09-24
/ NUMBER OF SEQ ID NOS: 15112
/ SEQ ID NO 156
/ LENGTH: 400
/ TYPE: DNA
/ ORGANISM: Bos taurus
/ OTHER INFORMATION: Clone ID: 01-LIB34-020-Q1-E2-A9
/ US-09-960-352-156

Query Match
Best Local Similarity 9.3%; Score 297; DB 3; Length 400;
Matches 343; Conservative 0; Mismatches 45; Indels 4; Gaps 2;
```

QY	1587	TGAGGGCCGGAGGAAGAAATTCCCAACGTATTAAGCGTATTCATAGCTAGGAAGTCT	1648
Db	9	TGGGGCCCGAGGAAGAAATTCCCACTGACAAAGCATCTTCATGCTTAAGAAAGTTC	68
QY	1647	TACCAACGAGCGAATATCTGAAGGATCTCGAAGTTATCACTTCGTGGTTCAGAGCAC	1706
Db	69	CACCACTGAGAGGACATATCTGAAGGATCTTGAAGTATCACTTCGTGGTTCAGAGCAC	128
QY	1707	AGTAGCAAAAGAGAGCGCCATGCGGGAAGCACTGAAAAGTCTCATATTTCCGAATTTTGA	1766
Db	129	AGTAGCAAAAGAGACTTCATGCCCCGAACCTTGAAGTCTCATATTTCCGAATTTTGA	188
QY	1767	ACCTTTGCAAAATTTCTATCAATTTTCTCAAGGAATTGAGCAAGACTTGCCTGTG	1826
Db	189	ACCTTTGCAAAATTTCTACACATATTTTCTCAAGGACATGAGCAAGACTTGCCTGTG	248
QY	1827	GGAAGGCGCGTCAAAATGCCCAATCA--GAGATTACCAAGAAATCGCGATGTCTATGCT	1883
Db	249	GGAAGGCGCGTCAAAATGCCCACTAGAGGAGATTACATAGAAATCGAGATGTATGTCT	308
QY	1884	GAAAGACATTAGGGCATGAAGCACCTTG-GGGCTCACCTGTGGAAGCACAGCGAGCCT	1942
Db	309	GAAAGACATTACGGGATGAAGCACTGAGCTGCTCACTTGTGGAAGCACAGTAGGCCC	368
QY	1943	TGGAGGCCCTGGAAGATGGAATCAAGACTCC	1974
Db	369	TGGAGGCCCTGGAACATCGGCATCTCTGGGCCCC	400

```

RESULT 14
US-10-062-674-58
: Sequence 58, Application US/10062674
: Publication No. US20040005559A1
: GENERAL INFORMATION:
: APPLICANT: Loring, Jeanne F.; Kaser, Matthew R.
: TITLE OF INVENTION: MARKERS OF NEURONAL DIFFERENTIATION AND MORPHOGENESIS
: FILE REFERENCE: Pa-0026-1 CIP
: CURRENT APPLICATION NUMBER: US/10/062,674
: CURRENT FILING DATE: 2002-01-30
: PRIOR APPLICATION NUMBER: US 09/625,102
: PRIOR FILING DATE: 2000-07-24
: NUMBER OF SEQ ID NOS: 2217
: SOFTWARE: PERL Program
: SEQ ID NO 58
: LENGTH: 437
: TYPE: DNA
: ORGANISM: Mus musculus
: FEATURE:
: NAME/KEY: miac feature
: OTHER INFORMATION: GenBank ID No. US20040005559A1 92305969
: US-10-062-674-58

```

Query Match	8.8%	Score 275.8	DB 7	Length 437
Beet Local	Similarly	Pred. No. 6.1e-72		
Matches 298	Conservative	0	Mismatches 37	Indels 0
				Gaps 0

Qy	2800	CAGTTGTCGGAAACCTGCTGAGGAAATTCAAAAACAGCAACGGGTGGCAGAACTGTGG	2855
Db	1	CAGCTCTCTGGGAACCTGCTGAGAAAGTTCAAGAACAGCAATGGCTGGCAAAAGCTGTGG	60
Qy	2860	GTGCTGTTCACAAACTTCTGCCTGTTCTTTCACAAATCACAACCAAGCATCATCCCTT	291.9
Db	61	GTGATATTCAACCACTTCTGCTGTTCTTTCACAAATCACAACCAAGCATCATCCCTT	120
Qy	2920	GCCAGCCTGCTCTGCTCGGCTACTCGCTCAACATCCCTCTGAGTCCGAGAACATCCAG	297.9
Db	121	GCCAGCCTGCTCTGCTCGGCTACTCCCTCAACATCCCTCTGAGTCCGAGAACATCCAG	180
Qy	2980	AAAGACATACGTTTCAAGCTGCACTTCAAAAGTCCCAAGCTACTTCAAGGGCGGAAGC	303.9
Db	181	AAAGACATACGTTTCAAGCTGCACTTCAAAAGTCCCAAGCTACTTCAAGGGCGGAAGC	240

QY 3040 GAGTACACGTCGAAAGGTGATGAAAGTATCCGACAGTGCACACAGCTTCCTCGGA 3059
DB 241 GAATTCACATTTGAAAGGTGATGAAAGTATCCGAAAGTGCACACAGTACGCTTCCTCG 300
QY 3100 CCCCACGTTGAGCCACAAAGAGTCTCTTGTA 3134
DB 301 GCCCACATCTTGAGTGCACAAAGAAATCTCACCTGTA 335

```

RESULT 15
US-10-450-763-565/c
; Sequence 565, Application US/10450763
; Publication No. US20050196754A1
GENERAL INFORMATION:
APPLICANT: HySeq, Inc
TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
FILE REFERENCE: 790CIP3/US
CURRENT APPLICATION NUMBER: US/10/450,763
CURRENT FILING DATE: 2003-06-11
PRIOR APPLICATION NUMBER: PCT/US01/08631
PRIOR FILING DATE: 2001-03-30
PRIOR APPLICATION NUMBER: 09/540,217
PRIOR FILING DATE: 2000-03-31
PRIOR APPLICATION NUMBER: 09/649,167
PRIOR FILING DATE: 2000-08-23
NUMBER OF SEQ ID NOS: 60736
SOFTWARE: Custom
SEQ ID NO 565
LENGTH: 1824
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SIMILAR
LOCATION: (1)..(306)
OTHER INFORMATION: 99% homologous to Homo sapiens ilal sodium-dependent bile
OTHER INFORMATION: acid transporter,accession number Z54350,Smith-Waterman Score=522
US-10-450-763-565

```

Query Match	8.1%	Score 254	DB 10	Length 1824
Best Local Similarity	100.0%	Pred. No. 5.2e-65		
Matches 254	0	Mismatches 0	Indels 0	Gaps 0

OY	1162	AAGCTCAGCAGACACACAGCTTACATTTGGAAAGGTGCGAAATCTCCAGGGGGGCGAG	1221
Db	1623	CAGTCTACGACGACACCAAGCTTACATTTGGAAAGGTGCCAAATCTCCAGGGGGCGAG	1564
OY	1222	AGCTGCCGCGCAGGAAAGAAACCGAAGTTTCCGCGGGGAGCCGGGGTCGACCCGAGC	1281
Db	1563	AGTGCCTGCGCGAGGAAAGGAAACCGAAGTTTCCGCGGGGAGCCGGGGTCGACCCGAGC	1504
OY	1282	CCTGCCCGCGAGGAAAGCCCCCGGGGTAAACAAGAGGCGGACCGGAGCGGCTCGGGCGCC	1341
Db	1503	CCTTCGCGCGAGGAAAGACCCCCCGGGGTAAACAAGAGGCGGACCGGAGCGGCTCGGGCGCC	1444
OY	1342	ACGAGGAAGAGAGAGGTCTGTTAAGATTAAGACCCAGCAGAGTAAACCTCAAGCCCGG	1401
Db	1443	ACGAGGAAGAGAGAGGTCTGTTAAGATTAAGACCCAGCAGAGTAAACCTCAAGCCCGG	1384
OY	1402	CAGCCAAAGCAGCAGG	1415
Db	1383	CAGCCAAAGCAGCAGG	1370

Search completed: May 22, 2006, 11:24:25
Job time : 3603.17 secs

This Page Blank (uspto)

GenCore version 5.1.8
Copyright (c) 1993 - 2006 BioCeleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 22, 2006, 09:18:29 ; Search time 32.413 Seconds
(without alignments)
6824.209 Million cell updates/sec

Title: US-09-555-342b-1_COPY_49_3183

Sequence: 1 atgggggaatagacagag.....acaagagctctgtgtat 3135

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 224314 seqs, 35277956 residues

Total number of hits satisfying chosen parameters: 448628

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database:

Published Applications NA New:
1: /EMC_Celera_SIDS3/prodata/1/pubnna/US09_NEW_PUB.seq.*
2: /EMC_Celera_SIDS3/prodata/1/pubnna/US06_NEW_PUB.seq.*
3: /EMC_Celera_SIDS3/prodata/1/pubnna/US07_NEW_PUB.seq.*
4: /EMC_Celera_SIDS3/prodata/1/pubnna/US08_NEW_PUB.seq.*
5: /EMC_Celera_SIDS3/prodata/1/pubnna/PCT_NEW_PUB.seq.*
6: /EMC_Celera_SIDS3/prodata/1/pubnna/US10_NEW_PUB.seq.*
7: /EMC_Celera_SIDS3/prodata/1/pubnna/US11_NEW_PUB.seq.*
8: /EMC_Celera_SIDS3/prodata/1/pubnna/US60_NEW_PUB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	192.4	6.1	2867	6 US-10-511-937-344	Sequence 344, App
2	56.8	1.8	3953	6 US-10-473-173-56	Sequence 56, App
3	41.8	1.3	6466	6 US-10-488-619-2736	Sequence 2736, App
4	41.4	1.3	37426	6 US-10-473-173-32	Sequence 32, App
5	40.4	1.3	3708	7 US-11-325-764-1	Sequence 1, App1
6	38.4	1.2	2367	6 US-10-505-928-744	Sequence 744, App
7	36.4	1.2	2711	7 US-11-145-307A-189	Sequence 189, App
8	35.6	1.1	1333	6 US-10-505-928-380	Sequence 380, App
9	35.4	1.1	4773	6 US-10-511-455-1	Sequence 1, App1
10	34.4	1.1	909	6 US-10-473-173-50	Sequence 50, App
11	34.2	1.1	4450	6 US-10-505-928-275	Sequence 275, App
12	34	1.1	4234	7 US-11-312-958-25	Sequence 25, App1
13	34	1.1	4573	6 US-10-505-928-409	Sequence 409, App
14	34	1.1	4585	6 US-10-505-928-47	Sequence 47, App1
15	33.8	1.1	555	7 US-11-217-529-17434	Sequence 17434, App
16	33.8	1.1	2121	6 US-10-196-749-355	Sequence 355, App
17	33.8	1.1	2121	7 US-11-101-316-117	Sequence 117, App
18	33.8	1.1	4740	6 US-10-511-937-380	Sequence 380, App
19	33.6	1.1	8833	6 US-10-505-928-785	Sequence 785, App
20	33.6	1.1	885	7 US-11-217-529-17330	Sequence 17330, App
21	33.4	1.1	5411	6 US-10-504-120-5	Sequence 5, App1
22	33.4	1.1	5543	6 US-10-504-120-6	Sequence 6, App1
23	33	1.1	2294	6 US-10-511-937-513	Sequence 513, App
24	32.6	1.0	819	7 US-11-024-544A-119	Sequence 119, App
25	32.6	1.0	819	7 US-11-024-545-47	Sequence 47, App1

26	32.6	1.0	819	7 US-11-185-301-35	Sequence 35, App1
27	32.6	1.0	819	7 US-11-190-750-102	Sequence 102, App
28	32.6	1.0	819	7 US-11-251-466-21	Sequence 21, App1
29	32.6	1.0	819	7 US-11-254-173-35	Sequence 35, App1
30	32.6	1.0	819	7 US-11-254-173-35	Sequence 35, App1
31	32.6	1.0	12649	7 US-11-264-784-22	Sequence 22, App1
32	32.6	1.0	12649	7 US-11-024-544A-110	Sequence 110, App
33	32.6	1.0	12649	7 US-11-024-545-38	Sequence 38, App1
34	32.6	1.0	12649	7 US-11-185-301-27	Sequence 27, App1
35	32.6	1.0	12649	7 US-11-190-750-93	Sequence 93, App1
36	32.6	1.0	12649	7 US-11-251-466-12	Sequence 12, App1
37	32.6	1.0	12649	7 US-11-254-173-26	Sequence 26, App1
38	32.6	1.0	12649	7 US-11-264-784-114	Sequence 114, App
39	32.4	1.0	2811	6 US-10-511-937-648	Sequence 648, App
40	32.4	1.0	4050	6 US-10-511-937-663	Sequence 663, App
41	32.4	1.0	4593	6 US-10-524-021-8	Sequence 8, App1
42	32.2	1.0	550	6 US-10-505-928-701	Sequence 701, App
43	32.2	1.0	627	6 US-10-488-619-2469	Sequence 2469, App
44	32.2	1.0	1086	7 US-10-488-619-2468	Sequence 2468, App
45	32.2	1.0	2040	7 US-11-217-529-1630	Sequence 1630, App
					Sequence 5680, App

ALIGNMENTS

RESULT 1
US-10-511-937-344
Sequence 344, Application US/10511937
GENERAL INFORMATION:
Publication No. US2006008836A1
APPLICANT: EXPRESSION DIAGNOSTICS, INC.
APPLICANT: Wohlgemuth, Jay
APPLICANT: Fry, Kirk
APPLICANT: Woodward, Robert
APPLICANT: Ly, Ngoc
APPLICANT: Prentice, James
APPLICANT: Morris, Macdonald
APPLICANT: Rosenberg, Steven
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
TITLE OF INVENTION: AND MONITORING TRANSPLANT REJECTION
FILE REFERENCE: 506612000104
CURRENT FILING DATE: US/10/511, 937
CURRENT FILING DATE: 2004-10-19
PRIOR APPLICATION NUMBER: PCT/US2003/012946
PRIOR FILING DATE: 2003-04-24
PRIOR APPLICATION NUMBER: US 10/131, 831
PRIOR FILING DATE: 2002-04-24
PRIOR APPLICATION NUMBER: US 10/325, 899
PRIOR FILING DATE: 2002-12-20
NUMBER OF SEQ ID NOS: 3117
SOFTWARE: PatentIn version 3.2
SEQ ID NO 344
LENGTH: 2867
TYPE: DNA
ORGANISM: Homo sapiens
US-10-511-937-344

Query Match 6.1%; Score 192.4; DB 6; Length 2867;
Best Local Similarity 52.8%; Pred. No. 7.5e-43;
Matches 490; Conservative 0; Mismatches 426; Indels 12; Gaps 3;

QY 117 CGGTCATCAAAATACAGATGCTGATGACCCAGGAGCACTTGAGTCCACAAAG 176
DB 798 CATGACCTGAGAGGCTTCTTGTGATGACACAGTTTAAATGTTGTGGAGAAACA 857
QY 177 AGCTCTGGAGAGTCTGCTGATGACAGTTTGCAACCACTTCAACTCGTGAAGTGA 236
DB 858 TGTGTAAGGACAAAGATTCTTAAAGAGATGATGACATTCATCTTTGGAAAGCA 917
QY 237 CTAATTTGGCTTGAATTTCCGATCAACAAAAGATCAGCGTGTGCTGATCTCTAA 296
DB 918 CTAATTTGGCTTGAATTTGGATTAACCAACCTTGAAGACATGCTGATTCGCCCA 977

QY	297	1CCCATTTGTGAAACGATTTAGAGGCCAAAGACGTTGTGTTAGTTGTGTGTAATTT	356
Db	978	AGAAATAAAAAGCAGGTTCCGGG--TGTCCTTGGAAATTTTACATTAATGTAAGTT	1034
QY	357	CTTTCGCGCTGACCAACACAACTCCAGAAGAACTCACAAGATACCTGTTGGCCCTGCA	416
Db	1035	TTATTCACCTBACCAGCAGAGTTTAACGAAGACATTAACAAATTTATTATGTCTTCA	1094
QY	417	GGTGAAGCAGGACTTTGGCTCAAGCAGGTTGACGTGTATGACACCAAGCAGCTCTCTT	476
Db	1095	GCTTCGGCAGGACATAGTTGACAGGACGTCCTGGTCTCTTTGGCAACTTTCAGATTAAT	1154
QY	477	GATTTCAACATTTGTCATCTGAGATTTGGGGATTTTGAATGAAGCCTTGGACAGAGACA	536
Db	1155	AGTTCTTTAACCATTCACGCTGAACTGGAGACCTACGACCCAGAACTTCATAGGGGTGA	1214
QY	537	CTTAGCAAAAAATAATACATAC-----TCAGCAAGACGACATAGAGACAAATCGT	590
Db	1215	TTATGTTAGTGAATTTTAAACGTGCCCCGGAATACAGACCAAGAACTTGAAGABAAGTCA	1274
QY	591	GGAATTTACACATACCAACATTTGACAAACACACAGACGAATCAGATTTCCAGCTCTTGA	650
Db	1275	GGAATCGATTAAGTCATACAGGTCATGACCTCAGACTCAGGCTGACTTGGATTTCTTGA	1334
QY	651	GATTGCCCGTGGCTGAGATGTATGGAATCCGGTTGACACCCGGCCAAAGAAAGGAAAG	710
Db	1335	GAATGCCAAAAAGTTGCTATGTATGAGAGTTGATCTTCAATAAAGCAAAAGACTTGGAAAG	1394
QY	711	CACGAAATCATCTGGCCGTTGCCAACACGGGAATTTCTAGTGTTCAGAGTTTCACTTA	770
Db	1395	AGTAGATATCATCTTAGGTGTCTGCTTAGTGGCCTTCTGGTTTCAAAAGATTAAGCTGAG	1454
QY	771	GATCAATGCTTTCAACTGGGCCCAAGGTGCGGAAGCTGAGCTTCAAGAGAAACGCTTCT	830
Db	1455	AATTAACGGCTTCCCTTGGCCCAAGTGTGTAAGATTTCTTTAAACGTAAGACTTTT	1514
QY	831	CATCAAGCTCCGGCCAGATGCCAATAGTGCCTAACAGATACCTTGGATTTCTGTATGCG	890
Db	1515	CATCAAGATTTGGCCTTGAGAGCAAGACAGATTAAGTAAGTACATTCGGATTTCAAACTTCC	1574
QY	891	CAGTCGGGATTTTCGCAAGTCCCTCTGGAAATTTGTGTGAACATCATGACCTTCTTAG	950
Db	1575	CAGTTTACCGAGACGCTAAGAAATTTATGAAAGTCTGTGTAGAACTACACGCTTTTTCAG	1634
QY	951	ACTTTTGAAGAGCCCAACCAAAAGCCCAAGCCGCTCTTTTAAACGGGGGTCAATCATT	1010
Db	1635	A---TTGACATCTACAGACACCATTTCCAAAGCAAAATTTCTTGGGCTAGGATCCAAAT	1691
QY	1011	TCGGTTCAGTGTGGGATCTCAGAAAGCAG	1038
Db	1692	TCGATACAGTGGCGGACTCAAGCTCAG	1719

```

RESULT 2
US-10-473-173-56
Sequence 56, Application US/10473173
Publication No. US2006008823A1
GENERAL INFORMATION:
APPLICANT: VAN ANBEL INSTITUTE
TITLE OF INVENTION: Microarray Gene Expression Profiling in Clear Cell Renal Cell
FILE OF INVENTION: Carcinoma: Prognosis and Drug Target Identification
FILE REFERENCE: 38345-170094
CURRENT APPLICATION NUMBER: US/10/473,173
CURRENT FILING DATE: 2003-09-29
PRIOR APPLICATION NUMBER: US 60/279,411
PRIOR FILING DATE: 2001-03-29
NUMBER OF SEQ ID NOS: 498
SOFTWARE: PatentIn version 3.2
SEQ ID NO 56
LENGTH: 3953
TYPE: DNA
ORGANISM: Homo sapiens
US-10-473-173-56

```

Query Match	1.8%	Score 56.8;	DB 6;	Length 3953;
Best Local Similarity	46.6%	Pred. No. 6.6e-06;		
Matches 218; Conservative	0;	Mismatches 247;	Indels 3;	Gaps 1;

QY	130	ATCCAGTGTGGATGACACCCAGGAGCAATTGAATTCACAAAGAGCTCTGGGAAG	189
Db	284	ATCCACTGCTTAAATACGAGTTTGTGAGTTACCTGTCCGTGGAGAGCACTGGCCAG	343
QY	190	GTCCTGTGGATGCAGTTTGCACACACTCAACTCTGTGAAAGGTGACTAATTTTGGCTC	249
Db	344	GAAGCCTCGAGGCCGTGGCCACAGAGGCTGAGCTGGGGAGGTCACTTAACTTACGCTTC	403
QY	250	GAGTTTCCTGATATCAAAAAGATCAAGGTGTGGTGTGATCTCTTAAACCATTTGTGAAA	309
Db	404	TGTACTACAAACAAGCAAAATCAGCCCGGTGGGTAAATTGGAAAAACCTTTGAAGAG	463
QY	310	CAGAT--TAGAAGGCCCAAAGCAAGTGTGTTTGAATTTGTGTGAATTCCTTCCGCT	366
Db	464	CAGCTGTATAAATATGATTTGGAACCTACCGTCTAATTTTGGAGTGGTGTAAATGTGCT	523
QY	367	GACCAACAACACTCCAAAGAACTCAACAAGTACTGTTCCGGCTGCAGGTGAAGCA	426
Db	524	TCAGTTTCTCAGCTGCAGCAGAGATTAACAAGTATCAATTAATTCGCCAATTGAAGAA	583
QY	427	GACTGGCTCAAGGCAAGTGTGAGGTGAATATGACACCAAGCCGAGCTCTTGATTTACAC	486
Db	584	GATATCTTGGAGGAGAGTATTCCTTGTAACTTGAACAACAAGCAATTCAGCTAGCAGCTTA	643
QY	487	ATTGTGCATCTGAGATTGGGATTTTGTATGAACCTTGAACAGAGCACTTAGACAAA	546
Db	644	GCTGTTCAACGGATTTTGGTGACTTGTATCAGTATGAATCCAGAGCTTCTTCAGAAA	703
QY	547	AATTAATACATACCTCAACAAGCACTTGAAGAGCAAAATTCGTGAA	594
Db	704	TTTGCCTTGTTCCTGTGGAGTGGTAAACAAGTAAAAAAGTATGGAA	751

```

RESULT 3
US-10-488-619-2736
; Sequence 2736, Application US/10488619
; Publication NO. US20060099578A1
; GENERAL INFORMATION:
; APPLICANT: Greenlee, Winner and Sullivan, P.C.
; TITLE OF INVENTION: Human Mitochondrial DNA Polymorphisms, Haplogroups, Associations w
; TITLE OF INVENTION: Physiological Conditions, And Genotyping Arrays
; FILE REFERENCE: 98-01 WO
; CURRENT APPLICATION NUMBER: US/10/488,619
; CURRENT FILING DATE: 2004-03-01
; NUMBER OF SEQ ID NOS: 3040
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2736
;
; LENGTH: 646
;
; TYPE: DNA
;
; ORGANISM: Mus musculus
; US-10-488-619-2736

```

Query Match	1.34	Score 41.8	DB 6	Length 646
Best Local Similarity	49.84	Pred. No. 0.028		
Matches 106	Conservative 0	Mismatches 107	Indels 0	Gaps 0
Qy	2007	GCTGCAGAAAGTGTGTACCTACCGCTCAACACCTTCTCTCGGCGCATCGACCGGCT	2066	
Db	73	GCTGCAGAAAGTATATTGACATCTGCTGAGATGGTTTCTGTGACCCCGGTGCAGAAAT	132	
Qy	2067	CATGACCTACAAAGCATGGTCTGTGAGCGGCTGTGCAACACCAACCGCTGAAGCCAGCCGA	2128	
Db	133	CTGGAAGTATCCCTCCAGCTGGGGGAGCTGCTCAATGACACACACCCCGACAGGGA	192	
Qy	2127	CTTCAGGAGCTGCCGAGCCGCTTTGSCAGATGACGAGATGTGTGGCACAGCTTCACGG	2186	
Db	193	CTTTAAGAAATGTTAAGACTGCTTTGACGCGCATGAAATATGTGGCCAGCTCATCAACGA	252	

Qy 2187 TACGATGATCAAGTGAATTTCCAGAGCT 2219
Db 253 ACGGAACGAAGCTTGAATTCAGACAGAT 285

RESULT 4

US-10-473-173-32
; Sequence 32, Application US/10473173
; Publication No. US2006008823A1
; GENERAL INFORMATION:
; APPLICANT: VAN ANDEL INSTITUTE
; TITLE OF INVENTION: Microarray Gene Expression Profiling in Clear Cell Renal Cell
; TITLE OF INVENTION: Carcinoma: Prognosis and Drug Target Identification
; FILE REFERENCE: 38345-170094
; CURRENT APPLICATION NUMBER: US/10/473,173
; CURRENT FILING DATE: 2003-09-29
; PRIOR APPLICATION NUMBER: US 60/279,411
; PRIOR FILING DATE: 2001-03-29
; NUMBER OF SEQ ID NOS: 498
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 32
; LENGTH: 37426
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-473-173-32

Query Match 1.3%; Score 41.4; DB 6; Length 37426;
Best Local Similarity 54.2%; Pred. No. 0.37;
Matches 84; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

Qy 205 GTTGCACACCTCACTCACTGAGAGTGATATTTGCGCTCGAGTTTCTGATCAC 264
Db 11233 GTGTGACGCGACTGGGATCATAGAGTTGACTATTTGACTGCTTTACGGGTAGC 11292
Qy 265 AAAAGATCAAGGTGGCTGATCTCTTAAACCCATTGTGAACAGATTAGAAGGCCA 324
Db 11293 AAAGTGAAGTTTATGGCTTAAACCTGAAACCGGATCTCCAGCAGATGATGGCTTA 11352
Qy 325 AAGCAGCTTCTTTAAGTTTGGTGAATTTCTT 359
Db 11353 GCCCTTACAGGCTTAAACTTAGATCAAGTTCTT 11387

RESULT 5

US-11-325-764-1
; Sequence 1, Application US/11325764
; Publication No. US20060099640A1
; GENERAL INFORMATION:
; APPLICANT: Jegla, Timothy James
; APPLICANT: Witzel, Julie Dickson
; APPLICANT: ICGen, Inc.
; TITLE OF INVENTION: S102 and S104, Novel Potassium Channel Proteins from
; FILE REFERENCE: 018512-006810US
; CURRENT APPLICATION NUMBER: US/11/325,764
; CURRENT FILING DATE: 2006-01-04
; PRIOR APPLICATION NUMBER: US/09/921,159
; PRIOR FILING DATE: 2001-08-01
; PRIOR APPLICATION NUMBER: US 60/249,112
; PRIOR FILING DATE: 2000-11-15
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 3708
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(3708)
; OTHER INFORMATION: human S102 potassium channel alpha subunit
US-11-325-764-1

Query Match 1.3%; Score 40.4; DB 7; Length 3708;

Best Local Similarity 58.2%; Pred. No. 0.18;
Matches 71; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

Qy 2046 CTTGCGGCACCTGACCGGCTCATGCAAGAGTCTTGGAGCGCTGTGAAACA 2105
Db 1917 CCGCTGCGCGCTGACAGATCATCGCTTCATGATGTGGCCATCGACCTGAGGGACAGA 1976
Qy 2106 CCAACCGCGGACGACCGGACCTTCAAGGATCTGCCAGCCGCTTTGGCAGATTCACGA 2165
Db 1977 GCAACCGGCTTACGCAAGCGCGGTGGGGGAGCAGCAAGCTGACCTGCCACGGA 2036
Qy 2166 GA 2167
Db 2037 GA 2038

RESULT 6

US-10-505-928-744
; Sequence 744, Application US/10505928
; Publication No. US20060088532A1
; GENERAL INFORMATION:
; APPLICANT: Ludwig Institute for Cancer Research et al.
; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
; FILE REFERENCE: 28967/39178
; CURRENT APPLICATION NUMBER: US/10/505,928
; CURRENT FILING DATE: 2004-08-27
; PRIOR APPLICATION NUMBER: US 60/363,019
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: PatentIn 3.2
; SEQ ID NO 744
; LENGTH: 2367
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-505-928-744

Query Match 1.2%; Score 38.4; DB 6; Length 2367;
Best Local Similarity 46.3%; Pred. No. 0.5; 146; Indels 0; Gaps 0;
Matches 126; Conservative 0; Mismatches 146; Indels 0; Gaps 0;

Qy 1867 ATGCGGATGTCATGCTGAAGACATTCAGGGGATGAGCACCTGGCGCTCACCTGTG 1926
Db 448 ATCTCGAGCTCCGGAAGATGATCTCAAGGGCTCCAGCACCTCAACGCCCTGCTG 507
Qy 1927 AAGCAGCGGAGGCTTGAAGGCTTGAAGATGAATGAAGCTCCGCGGCTGAG 1986
Db 508 GTGAACAAAGATCTCAAGATCCATGAGAAAGGCTTCAAGCCCACTGCGGAAAGCTGCA 567
Qy 1987 AACTCTGAGAGACTTGAAGCTGCAAGAGGTGTACTACCGCTCAACACCTTCTC 2046
Db 568 AGCTTACATCTTCAAGAACCACTGTGGAGATCCGCGCCCACTTACCCAGCTCTG 627
Qy 2047 CTGCGGCACATGACCGGCTCATGCACTCAAGAGAGTCTTGGAGCGGCTGCAACAC 2106
Db 628 GTGAGCTCGGATGACAGCAACACCATCCGAGAGTGTCCAGAGGATGTTCACAGGG 687
Qy 2107 CACCGCGGAGCGGCGGACCTTCAGGACTG 2138
Db 688 CTCGGGACATGAATGATGAGATGGGCGG 719

RESULT 7

US-11-145-307A-189/c
; Sequence 189, Application US/11145307A
; Publication No. US20060094035A1
; GENERAL INFORMATION:
; APPLICANT: Actarus Bioscience, Inc.
; APPLICANT: Erlander, Mark G.
; TITLE OF INVENTION: Identification of Tumors
; FILE REFERENCE: 022041-002020US
; CURRENT APPLICATION NUMBER: US/11/145,307A
; CURRENT FILING DATE: 2005-06-03

PRIOR APPLICATION NUMBER: US 60/577,084
PRIOR FILING DATE: 2004-06-04
NUMBER OF SEQ ID NOS: 268
SOFTWARE: PatentIn version 3.1
SEQ ID NO 189
LENGTH: 2711
TYPE: DNA
ORGANISM: Homo sapiens
US-11-145-307A-189

Query Match 1.2%; Score 36.4; DB 7; Length 2711;
Best Local Similarity 56.8%; Pred. No. 1.9;
Matches 67; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 147 CACCCAGAGGAGCTTGAAGTTCACAAAGAGCTCTGGGAGAGGTGCTGATGATCAGT 206
DB 500 CAGCCAGAGGAGGCTCTCGGATTCCTCAGGCAACATGGTAGGTTCTGGCGTTTATG 441
QY 207 TTGCAACACCTCAACCTCGTGAAGTGAATATTTTGGCTTGAAGTTCTTGATCAG 264
DB 440 CCCCATCTTCTCCACACCTTGGGGGTGGCCACTTGGGCTTGGAGCCCTATCAG 383

RESULT 8

US-10-505-928-380/C
Sequence 380, Application US/10505928
Publication No. US20060088532A1
GENERAL INFORMATION:
APPLICANT: Ludwig Institute for Cancer Research et al.
TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
FILE REFERENCE: 28967/39178
CURRENT APPLICATION NUMBER: US/10/505,928
CURRENT FILING DATE: 2004-08-27
PRIOR APPLICATION NUMBER: US 60/363,019
PRIOR FILING DATE: 2002-03-07
NUMBER OF SEQ ID NOS: 866
SOFTWARE: PatentIn 3.2
SEQ ID NO 380
LENGTH: 1323
TYPE: DNA
ORGANISM: Homo sapiens
US-10-505-928-380

Query Match 1.1%; Score 35.6; DB 6; Length 1323;
Best Local Similarity 42.8%; Pred. No. 2.1;
Matches 179; Conservative 0; Mismatches 239; Indels 0; Gaps 0;

QY 2274 GGAATTCATCCGTCCTGGGAGCCTCAGCAAGCTCTGGGGAAGGGCTCCAGCAGCTG 2333
DB 957 GGGGCTCATGGCAGGCGGGTTTCGAGCACTGTCTGGGCAATGAAAGAGGGGCGAG 898
QY 2334 GTTCTTCTGTTCAAGAGCTCTGCTATACAGCCGGGGGCTGACGGCTCCAAATCA 2393
DB 897 CCAACACAGCTGGGCAACCAATGATCCCAAGAGCTGAGCCATATCTCCAGCTCGGA 838
QY 2394 GTTTAAAGTCAAGGAGCTCCGCTCTATATGAGCATGAGATGAGAGAGCAAGACGA 2453
DB 837 GTCCCGGGAGCGCTGTGGGCGGCTCTGCGCCGTGTAGACGTGCAACAGGTTGGCAC 778
QY 2454 GTGGGGGGTGGCCCACTGCTGACCTCCGGGGGCGAGCGGCACTCATCATGTGGCCGC 2513
DB 777 GGTGACCGTGTTCAGAGAGAGAGAGCCGACGAGAGGCGCCCAAGATGAGAAAG 718
QY 2514 CAGTTCTCGTCCGAGATGAGAGAGTGGGTGAGGACATCCAGATGAGCCATTGACCTG 2573
DB 717 CAGCCCGAAGGCTCTCCCGGACTCGGCGCCAGCGTACAGAGACCAAGAGCCCGG 658
QY 2574 GGAAGAGAGAGAGCCCGCCCTGAGTTCCTGGCAGAGAGCCCTGACCAAGATC 2633
DB 657 GATTTGACAGGTGTACGAGCCAGCGCCAGAGGGGCAAGAGCCAGGCGGAGCGGCG 598
QY 2634 CCTGTAAAGCAGCGCGGCTGACAGAGATGAGAGATGACCTTGAAGGCGCTCGGCG 2691

DB 597 CGCCACACAGCCCAAGGTGGCCCAAGCGGGCGCTGCGAGGAGCCGCGGGCGC 540

RESULT 9

US-10-511-455-1
Sequence 1, Application US/10511455
Publication No. US20060088835A1
GENERAL INFORMATION:
APPLICANT: Pickard, Benjamin Simon
APPLICANT: Blackwood, Douglas
APPLICANT: Porteous, David
APPLICANT: Muir, Walter John
APPLICANT: Muir, Walter John
APPLICANT: Moys, Ole
APPLICANT: Ewald, Henrik Lynke
TITLE OF INVENTION: SCHIZOPHRENIA ASSOCIATED GENES
FILE REFERENCE: 9013.63
CURRENT APPLICATION NUMBER: US/10/511,455
PRIOR FILING DATE: 2004-10-05
PRIOR APPLICATION NUMBER: PCT/GB03/001543
PRIOR FILING DATE: 2003-04-07
PRIOR APPLICATION NUMBER: GB0207902.8
PRIOR FILING DATE: 2002-04-05
PRIOR APPLICATION NUMBER: GB0207904.4
PRIOR FILING DATE: 2002-04-05
PRIOR APPLICATION NUMBER: GB0207900.2
PRIOR FILING DATE: 2002-04-05
PRIOR APPLICATION NUMBER: GB0207901.0
PRIOR FILING DATE: 2002-04-05
PRIOR APPLICATION NUMBER: GB0227734.1
PRIOR FILING DATE: 2002-11-28
NUMBER OF SEQ ID NOS: 94
SOFTWARE: PatentIn version 3.1
SEQ ID NO 1
LENGTH: 4773
TYPE: DNA
ORGANISM: Homo sapiens
US-10-511-455-1

Query Match 1.1%; Score 35.4; DB 6; Length 4773;
Best Local Similarity 44.5%; Pred. No. 4.9;
Matches 141; Conservative 0; Mismatches 176; Indels 0; Gaps 0;

QY 2411 AGCTCCGCTCTTANGCATGACGATTTGAGAGAGGAAAGCAGATGGGGGTGCCCACT 2470
DB 2219 AACTCCAGCTGATGAGGCTGTGATGATGATGACAGAAAGCACTTCTGATGACCTGC 2278
QY 2471 GCTTACCTCCGGGGCCAGCGGCACTCATCATGTGGCCGCGCTTCTGGTCCAGA 2530
DB 2279 ACATGACATGTGTGAGAGAGGACGACACAGGCCATTCATTCACAGCTAGGCTGC 2338
QY 2531 TGAAGAAGTGGTGTGAGACATCCAGATGGCCATTGACCTGGCGGAGAAAGCAGCAGCC 2590
DB 2339 AGCAGAAAGAGACGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2398
QY 2591 CCGCCCTGAGTCTCTGCGCAGCAGCCCTCCGACCAAGTCCCTGATGAAGCCACG 2650
DB 2399 AGCAGAAAGAGACGAGCGGTGTGGGCGGAGCCGAGAGGAGCCTCGTATATGACGAG 2458
QY 2651 CGGCTGACCAAGAGTCAAGATGACCTGAGGCGCTCCGCGCACTGCGGAGGCGCAGG 2710
DB 2459 CGGAGCAAGAGAAATGAGCGAGCGAGCGACCGCATCTCCACCGCTGGGCGGAGCA 2518
QY 2711 CCGCCAGCCGCGGCAAC 2727
DB 2519 GGAAGCTCACCTGACG 2535

RESULT 10

US-10-473-173-50
Sequence 50, Application US/10473173
Publication No. US20060088823A1
GENERAL INFORMATION:
APPLICANT: VAN ANDEL INSTITUTE

```

; TITLE OF INVENTION: Microarray Gene Expression Profiling in Clear Cell Renal Cell
; FILE REFERENCE: 38345-170094
; CURRENT APPLICATION NUMBER: US/10/473,173
; CURRENT FILING DATE: 2003-09-29
; PRIOR APPLICATION NUMBER: US 60/279,411
; PRIOR FILING DATE: 2001-03-29
; NUMBER OF SEQ ID NOS: 498
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 50
; LENGTH: 909
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-473-173-50

Query Match      1.1%; Score 34.4; DB 6; Length 909;
Best Local Similarity 53.8%; Pred. No. 3.5;
Matches 71; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

QY      1252 TCCGCGCGGAGCGCGGGGTGCGACCCGAGCCCTGCGCGGAGGAGGAGCCCGCGGGTAAAC 1311
DB      670 TAGCGCGCGGAGCGCGGGGTGCGACCCGAGCCCGAGGAGGAGGAGGAGCGCGGGGCGAGC 729
QY      1312 AAGCAGCGGAGCGGAGCGCGCTCGCGCGCCGACGAGGAGGAGGAGGAGGAGTCTTAAGAT 1371
DB      730 CGGAGCGCGGGCGCGCGCGCGGACCCCTGCGCGAGGAGTGGCGCCCTTGGGGCG 789
QY      1372 AGGACCCGACGAG 1383
DB      790 CCGCACCCTCCAG 801
```

RESULT 11

```

US-10-505-928-275
; Sequence 275, Application US/10505928
; Publication No. US20060088532A1
; GENERAL INFORMATION:
; APPLICANT: Ludwig Institute for Cancer Research et al.
; FILE REFERENCE: 28967/39178
; CURRENT APPLICATION NUMBER: US/10/505,928
; CURRENT FILING DATE: 2004-08-27
; PRIOR APPLICATION NUMBER: US 60/363,019
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: PatentIn 3.2
; SEQ ID NO 275
; LENGTH: 4450
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-505-928-275
```

```

Query Match      1.1%; Score 34.2; DB 6; Length 4450;
Best Local Similarity 49.7%; Pred. No. 9.9;
Matches 87; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY      2058 GCACCGGCTCATCATCAAGAGAGGCTCTGAGAGCGGCTGTGCAACACACCCGCGGAG 2117
DB      2010 GCACCTGCGCAAGAGAGAGTCTGTGCGTGCAGGCGCTGGAAGCCCTGCGCTCAAGCGAGAA 2069
QY      2118 CCAGCGCGGACTTGTGAGGAGTCCGAGCGGCTTTGGCAGAGATCATCCGAGATGGTGGACA 2177
DB      2070 CTGAGCTGAGCTCTGTGTGAACGTGAGCGGAGCTGCTGGAATGATGATGCTTGGTGGCGG 2129
QY      2178 GCTTCACGGTACATGATCAAGATGAGAAATTTCCAGAAAGCTGACGAACTCAAG 2232
DB      2130 AGCGCAGCGCCGACGATCGTGTGTGTCAGAAAGAGAGAGGCTGTGAGAGAAAG 2184
```

```

RESULT 12
US-11-312-958-25
; Sequence 25, Application US/11312958
; Publication No. US20060100152A1
```

```

; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Rosenfeld, Julie Beth
; APPLICANT: Siles-Santillan, Immaculada
; TITLE OF INVENTION: METHODS AND COMPOSITIONS IN TREATING
; TITLE OF INVENTION: PAIN AND PAINFUL DISORDERS USING 9949, 14230, 760, 62553,
; TITLE OF INVENTION: 12216, 17719, 41897, 47174, 33408, 10002, 16209, 314, 636,
; TITLE OF INVENTION: 27410, 33360, 619, 15985, 69112, 2158, 224, 615, 44373,
; TITLE OF INVENTION: 95431, 22245, 2387, 16658, 55054, 16314, 1675, 9569 OR
; TITLE OF INVENTION: 13424 MOLECULES
; FILE REFERENCE: MP102-027P1MONMIM
; CURRENT APPLICATION NUMBER: US/11/312,958
; CURRENT FILING DATE: 2005-12-20
; PRIOR APPLICATION NUMBER: US/10/369,022
; PRIOR FILING DATE: 2003-02-19
; PRIOR APPLICATION NUMBER: US 60/360,495
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/370,121
; PRIOR FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: US 60/373,010
; PRIOR FILING DATE: 2002-04-16
; PRIOR APPLICATION NUMBER: US 60/373,908
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: US 60/377,717
; PRIOR FILING DATE: 2002-05-03
; PRIOR APPLICATION NUMBER: US 60/379,949
; PRIOR FILING DATE: 2002-05-13
; PRIOR APPLICATION NUMBER: US 60/382,409
; PRIOR FILING DATE: 2002-05-21
; PRIOR APPLICATION NUMBER: US 60/385,280
; PRIOR FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US 60/386,879
; PRIOR FILING DATE: 2002-06-06
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 4234
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURES:
; NAME/KEY: CDS
; LOCATION: (863)...(2452)
US-11-312-958-25
```

```

Query Match      1.1%; Score 34; DB 7; Length 4234;
Best Local Similarity 53.8%; Pred. No. 11;
Matches 70; Conservative 0; Mismatches 60; Indels 0; Gaps 0;
```

```

QY      2240 TGATTGGCATTCATCTTTGTGTTTCCGGAAGGAGTTCACTCGTGGGAGGCTCA 2299
DB      1908 TGTTCGCACTTCTTCAAGCTCTTCCGCACTCCAAAGGGGTGCGAGATCTGGGCAAGACTT 1967
QY      2300 GCAAGCTTCGCGGAGGAGGCTCCAGAGGAGTGTCTTCTTCTTCAACGAGTCTGCG 2359
DB      1968 TCGAGGCTTCATGAGAGGAGCTGGGCTGTCTCTCTTCTTCTTCTTCAATGGGATCATCC 2027
QY      2360 TATACACGAG 2369
DB      2028 TCTTCTCCAG 2037
```

```

RESULT 13
US-10-505-928-409
; Sequence 409, Application US/10505928
; Publication No. US20060088532A1
; GENERAL INFORMATION:
; APPLICANT: Ludwig Institute for Cancer Research et al.
; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
; FILE REFERENCE: 28967/39178
; CURRENT APPLICATION NUMBER: US/10/505,928
; CURRENT FILING DATE: 2004-08-27
; PRIOR APPLICATION NUMBER: US 60/363,019
```

/ PRIOR FILING DATE: 2002-03-07
/ NUMBER OF SEQ ID NOS: 866
/ SOFTWARE: PatentIn 3.2
/ SEQ ID NO 409
/ LENGTH: 4573
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-505-928-409

Query Match 1.1%; Score 34; DB 6; Length 4573;
Best Local Similarity 59.2%; Pred. No. 11;
Matches 58; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

QY 1494 GTCTCCCAACTGAGCCCGACACCAAGCAGGCTCTCTCTGTATCAGCCCGCTGTGAA 1553
DB 437 GTGCGGATTTGGAGCAGCAGAGCTCCAGAGTTCTGCCCCACCATCTCTCAGCAGCTGGA 496

QY 1554 TGACCAAGGCTGCCCCGAGCAGACGATGAGATGAGG 1591
DB 497 TTCCCGGCGCTGCACCTCGGAGAACCAAGAAACGAGG 534

RESULT 14
US-10-505-928-47
/ Sequence 47, Application US/10505928
/ Publication No. US20060088532A1
/ GENERAL INFORMATION:
/ APPLICANT: Ludwig Institute for Cancer Research et al.
/ TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
/ FILE REFERENCE: 28967/39178
/ CURRENT APPLICATION NUMBER: US/10/505,928
/ PRIOR FILING DATE: 2004-08-27
/ PRIOR APPLICATION NUMBER: US 60/363,019
/ PRIOR FILING DATE: 2002-03-07
/ NUMBER OF SEQ ID NOS: 866
/ SOFTWARE: PatentIn 3.2
/ SEQ ID NO 47
/ LENGTH: 4585
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ OTHER INFORMATION: KIAA0062
US-10-505-928-47

Query Match 1.1%; Score 34; DB 6; Length 4585;
Best Local Similarity 59.2%; Pred. No. 11;
Matches 58; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

QY 1494 GTCTCCCAACTGAGCCCGACACCAAGCAGGCTCTCTCTGTATCAGCCCGCTGTGAA 1553
DB 449 GTGCGGATTTGGAGCAGCAGAGCTCCAGAGTTCTGCCCCACCATCTCTCAGCAGCTGGA 508

QY 1554 TGACCAAGGCTGCCCCGAGCAGACGATGAGATGAGG 1591
DB 509 TTCCCGGCGCTGCACCTCGGAGAACCAAGAAACGAGG 546

RESULT 15
US-11-217-529-174434
/ Sequence 174434, Application US/11217529
/ Publication No. US20060099612A1
/ GENERAL INFORMATION:
/ APPLICANT: SUNTORY LIMITED
/ APPLICANT: NAKAO, YOSHIHIRO
/ APPLICANT: NAKAMURA, NORIHIISA
/ APPLICANT: KODAMA, YUKIKO
/ APPLICANT: FUJIMURA, TOMOKO
/ APPLICANT: ASHIKARI, TOSHIHIKO
/ TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
/ FILE REFERENCE: S-38-285
/ CURRENT APPLICATION NUMBER: US/11/217,529
/ CURRENT FILING DATE: 2005-09-02
/ PRIOR APPLICATION NUMBER: US 10/932,182

/ PRIOR FILING DATE: 2004-09-02
/ NUMBER OF SEQ ID NOS: 197023
/ SOFTWARE: PatentIn version 3.3
/ SEQ ID NO 174434
/ LENGTH: 555
/ TYPE: DNA
/ ORGANISM: Saccharomyces pastorianus
US-11-217-529-174434

Query Match 1.1%; Score 33.8; DB 7; Length 555;
Best Local Similarity 48.2%; Pred. No. 3.9;
Matches 95; Conservative 0; Mismatches 102; Indels 0; Gaps 0;

QY 2532 GGAGAAAGTGGATTGAGACATCAGATGCGCATTTGACCTGCGGAGAAAGCAGCAGCCCC 2591
DB 246 GACACAGCTGTTCTTGACAAAGAAAGCCATAGATCTGACGAGAGCCCGACACAC 305

QY 2592 CCCCCCTGAGTTCTGGCCAGCAGCCCCCTGACAAAGTCCCTGATGAAGCCACCGC 2651
DB 306 AAACACTACGATCAATTAATTAACAACATGACAGCCCGTTGACGAGAGCCCGA 365

QY 2652 GGCTGACCAAGATGAGGATGACCTGAGCGCTCGGCAATGCGTGGAGCGCCAGGC 2711
DB 366 CAGCAGCATTAATTAATTAATCAAGAGAGAGTCCATCAATTTCTGTGCGCTGATTTGTC 425

QY 2712 CCCGACCGCGGCAACA 2728
DB 426 CAGCAGCAGTACCAACA 442

Search completed: May 22, 2006, 11:25:34
Job time : 32.413 secs

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: May 20, 2006, 01:28:11 ; Search time 14245 Seconds

(without alignments)

7036.685 Million cell updates/sec

Title: us-09-555-342b-2

Perfect score: 5463

Sequence: 1 MBEIQRPTGSRRLGAPENS.....SATSSASRPHVLSHKSLEYV 1045

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 6366136 seqs, 31973710525 residues

Total number of hits satisfying chosen parameters: 12732272

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame+ .p2n.model -DEV=xlh
-Q=/abs/ABSSWEB.spool/US09555342/runat.19052006.121454.13184/app.query.fasta_1
-DB=GenEmbl -QPMF=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPEXT=0 -LIST=45
-UNITS=bits -START=1 -END=1 -MATRIX=blousum62 -TRANS=human40.cdi -HOST=abs07
-DOCALLGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFM=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abs07
-USRR=US09555342.OCGN.1.1.5548.0/runat.19052006.121454.13184 -NCPU=6 -ICPU=3
-NO_MMP -NEG_SCORES=0 -WAIT -DGBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:.*
1: gb_env:.*
2: gb_pat:.*
3: gb_ph:.*
4: gb_dl:.*
5: gb_pr:.*
6: gb_ro:.*
7: gb_atg:.*
8: gb_gy:.*
9: gb_un:.*
10: gb_vl:.*
11: gb_ov:.*
12: gb_hcg:.*
13: gb_in:.*
14: gb_cm:.*
15: gb_ba:.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	5463	100.0	3442	2 CS247579 Sequence
2	5463	100.0	3442	5 AB008430 Homo sapi
3	5463	100.0	3546	5 BC041595 Homo sapi

4	5463	100.0	5028	2 CS130723	CS130723 Sequence
5	5459	99.9	4687	2 BD231200	BD231200 Human cyt
6	5431.5	99.4	4932	5 BC071592	BC071592 Homo sapi
7	4095.5	75.0	2937	2 CQ724059	CQ724059 Sequence
8	3963.5	72.6	3902	6 BC030329	BC030329 Mus muscu
9	2941.5	53.8	3997	2 CQ895396	CQ895396 Sequence
10	2941.5	53.8	3997	5 AB018336	AB018336 Homo sapi
11	2936.5	53.8	3997	2 CQ725044	CQ725044 Sequence
12	2924	53.5	3719	6 BC009153	BC009153 Mus muscu
13	2117	38.8	2632	6 BC004009	BC004009 Mus muscu
14	1879.5	34.4	4048	13 BT024186	BT024186 Drosophill
15	1784	32.7	2431	6 BC027077	BC027077 Mus muscu
16	1669	30.6	2596	5 BC021301	BC021301 Homo sapi
17	1539	28.2	2296	6 BC043327	BC043327 Mus muscu
18	1437	26.3	129380	11 AY739094	AY739094 Trkifugu
19	1150.5	21.1	3200	2 CQ842377	CQ842377 Sequence
20	1150.5	21.1	3200	5 AK125336	AK125336 Homo sapi
21	1075	19.7	2436	11 BC110755	BC110755 Xenopus 1
22	913.5	16.7	1747	2 CQ575283	CQ575283 Sequence
23	905.5	16.6	1230	2 CQ572985	CQ572985 Sequence
24	905.5	16.6	2538	13 AY060484	AY060484 Drosophill
25	891.5	16.3	4150	2 CQ575282	CQ575282 Sequence
26	891.5	16.3	42014	12 AC018327	AC018327 Drosophill
27	891.5	16.3	160712	13 AC009537	AC009537 Drosophill
28	891.5	16.3	164712	13 AC007589	AC007589 Drosophill
29	891.5	16.3	303434	13 AB003604	AB003604 Drosophill
30	850.5	15.6	2888	5 HSM602608	AL161984 Homo sapi
31	798.5	14.6	3474	4 BT008434	BT008434 Arabidops
32	798.5	14.6	3592	4 BT002491	BT002491 Arabidops
33	790	14.5	235700	12 AC130919	AC130919 Rattus no
34	790	14.5	283769	12 AC094777	AC094777 Rattus no
35	754.5	13.8	3275	2 CQ850384	CQ850384 Sequence
36	754.5	13.8	3275	5 AK127510	AK127510 Homo sapi
37	754	13.8	3594	11 CT025506	CT025506 Xenopus t
38	754	13.8	3760	11 BC063888	BC063888 Xenopus t
39	753	13.8	3488	11 BC060449	BC060449 Xenopus 1
40	752.5	13.8	3488	6 AB032828	AB032828 Rattus no
41	752	13.8	4051	6 AF152247	AF152247 Mus muscu
42	748.5	13.7	3309	2 CS032039	CS032039 Sequence
43	748.5	13.7	3309	2 CS032937	CS032937 Sequence
44	748.5	13.7	3309	2 CS040991	CS040991 Sequence
45	748.5	13.7	3309	2 CS041889	CS041889 Sequence

ALIGNMENTS

RESULT 1	CS247579	3442 bp	DNA	linear	PAT 09-JAN-2006
LOCUS	CS247579				
DEFINITION	Sequence 31 from Patent EP1612281.				
ACCESSION	CS247579				
VERSION	CS247579.1	GI:84661528			
KEYWORDS					
SOURCE					
ORGANISM	Homo sapiens (human)				
REFERENCE	Raponi, M.				
AUTHORS	Methods for assessing patients with acute myeloid leukemia				
TITLE	Patent: EP 1612281-A 31 04-JAN-2006;				
JOURNAL	Veridex, LLC (US)				
FEATURES	Location/Qualifiers				
source	1. .3442				
	/organism="Homo sapiens"				
	/mol_type="unassigned DNA"				
	/db_xref="taxon:9606"				

ORIGIN

Alignment Scores: 0 Length: 3442

Pred. No.: 5463.00 Matches: 1045

Score: 5463.00

Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	2	Gaps:	0
US-09-555-342B-2 (1-1045) x CS247579 (1-3442)			
QY	1	MetGlyGluIleGluGlnatgProthrProGlySerArgLeuGluValAProGluAenSer	20
DB	49	ATGGAGAAATATAGAGAGAGAGCCGACCCAGGATCACGCTGGGGGCCCCGAAATTCG	108
QY	21	GlyTlIeserThrLeuGluatrgGlyGlnLysProProProThrProSerGlyLysLeuVal	40
DB	109	GGGATCATGTAACCTTGAAACGTGACAGAAACCCGCCCAACACCTTCAGGAAATCTGTG	168
QY	41	SerTlIeLysTlIeGlnMetLeuAAspThrGlnGluValAphGluValAProGlnatrgA	60
DB	169	TTCATCAAAATCCAGATGCTGATGATGACACCCAGAGGCACTTTGAAAGTTCCACAAAGACT	228
QY	61	ProGlyLysValLeuLeuAAspAvalCysAAsnHsLeuAAsnLeuValGluGlyAspTyr	80
DB	229	CTGGGAGAGTGCTGGTGGATGATGACAGTTTGCAACCACTCAACCTCGTGGAGGTGACTAT	288
QY	81	PhGlyLysGluGluPheProAAspHsLysLysTlIeThrValTTrLeuAAspLeuLysPro	100
DB	289	TTTGGCTTCAGATTCTTGATCACAAAAGATCAACGGTGGCTGGATCTCTAAACC	348
QY	101	ILeValIleGlnIleArgArgProLysHsIleValIleLysPheValValLysPhePhe	120
DB	349	ATTGTGAACAGATTAGAGGCCAAAGCACTGTTGTTAAAGTTGGTGAATTCCTTT	408
QY	121	ProProAAspHsThrGlnLeuGlnGluLeuThrArgTyrLeuPheAAspLeuGlnVal	140
DB	409	CCGCTGACCAACACAACTCCAAGAAAGACTCAAGAGTACCTGTTCCGGCTGCAGGTG	468
QY	141	LysGlnAAspLeuAAspGlnGlyLeuThrCysAAsnAAspThrSerAlaAAspLeuIle	160
DB	469	AACAGAGACTTGGCTCAAGCAGGTGACGTGTAATGACACAGCGAGCTCTTGATT	528
QY	161	SerHsIleValGlnSerGluIleGlyAAspPheAAspGluAAspAAspArgGlnHsLeu	180
DB	529	TTCACACTTGTGCAATGTCGATTTGGGGAATTTGATGAGCCTTGACAGAGAGCACTTA	588
QY	181	AlaLysAAsnLysTyrIleProGlnIleAAspAAspLeuAAspLysIleValGlnPheHs	200
DB	589	GCAAAATATATATACCTCACTCAAGAAAGCACTAGAGGACAAATTCGGAAATTTAC	648
QY	201	HsAAsnHsIleGlyGlnThrProAAspLysAAspPheGlnLeuGlnIleAAsp	220
DB	649	CATTAACCACTTGACAAACACAGCAGATCAGATTTCCAGCTCTAGAGATTGCCCT	708
QY	221	ArgLeuGlnMetTyrGlyIleArgLeuHsIleProAlaLysAAspArgGluTyrLysIle	240
DB	709	CGGCTAGAGATGATAGAAATCCGGTTGCAACCGGCCAAGSACAGSAGSAGSAGATC	768
QY	241	AsnLeuAAspValAAsnThrGlyIleLeuValAphGlnGlyPheThrLysIleAAsnA	260
DB	769	AATCTGGCCGTTGCAACAGGAAATCTAGTGTTCAGAGGTTTCACTAAGATCAATGCC	828
QY	261	PhAAsnThrAlaLysValAAspLysLeuSerPheLysArgGlyAAspPheLeuIleLysLeu	280
DB	829	TTCAACTGGGCGCAAGGCGGAGCTGAGCTTCAAGAGGAAAGGCTTTTCATCAAGCTC	888
QY	281	ArgProAAspAAsnSerAlaTyrGlnAAspThrLeuGluPheLeuMetAlaSerAAsp	300
DB	889	CGGCGAGATGCCAATAGTGGCTACAGAGATACCTTGAATTCCTGATGGCAATCCGGAT	948
QY	301	PhCysLysSerPheTyrLysIleCysValGlnHsIleAlaPhePheArgLeuPheGlu	320
DB	949	TTCTGCAAGCTCTTTCGAAATATCTGTGTAACATCATCTCTTATGACTTTTGAAT	1008
QY	321	GluProLysProLysProLysProValLeuPheSerArgLysSerPheArgPheSer	340

DB	1009	GAGCCCAAAACAAAGCCCAAGCCCTCTCTTAAAGCCGGGGGTCAATCTTCGTTCAAGT	1068
QY	341	GlyArgThrGlnLysGlnValLeuAAspTyrValLysGluGlyValHsLysLysValGln	360
DB	1069	GGTCCGAGCTCAGAGAGAGTTCTCGACTATGTTAAAGAGAGACATTAAGAAAGTGCAG	1128
QY	361	PheGluAAspLysHsSerLysTlIeHsSerTlIeArgSerLeuAAspGlnProThrGlu	380
DB	1129	TTTGAAGAGAACACAGCAAGATTCATCTTACCGAGCTTGCTTCAACAGCTTACAGAA	1188
QY	381	LeuAAsnSerGluValLeuGlnGlnSerGlnGlnSerThrSerLeuThrPheGlyGlyGly	400
DB	1189	CTGAATTCGAAAGTGTGAGAGCATCTTCAAGAGCACAGCTTACATTTGGAGAGGT	1248
QY	401	AlaGluSerProGlyGlyGlnSerCysArgArgGlyLysGluProLysValSerAlaGly	420
DB	1249	GCGAATCTCCAGGGGGGCCAGAGCTGCGCGAGAGAAAGAACAGAGTTTCCCGCGGG	1308
QY	421	GluProGlySerHsIleProSerProAlaProArgArgSerProAlaLysAAspGlnA	440
DB	1309	GAGCCGGGCTCCACCCGAGCCCTGCGCGAGAGAGAGCCCGCGGTTAAACAGCAGGCG	1368
QY	441	AAspGlyAlaAlaSerAlaProThrGlnGlnGlnGlnValValLysAAspArgThrGln	460
DB	1369	GAGGAGCCGCTCGGCGCCACGAGAGAGAGAGAGAGGTTGTTAAGATAGAACCCAG	1428
QY	461	GlnSerLysProGlnProProGlnProSerThrGlySerLeuThrGlySerProHsLeu	480
DB	1429	CAGAGTAACTCAGGCCCGCGAGCCCAACAGAGTCTCTGATGCGAGTCCCTCACCTT	1488
QY	481	SerGluLeuSerValAAsnSerGlnGlyValAlaProAlaAAsnValThrLeuSerPro	500
DB	1489	TCCGAGCTCTCTGTGAATCCGAGGGGGAGTGGCCCTCCCAACGTGACTGTCTCCC	1548
QY	501	AAsnLeuSerProAAspThrLysGlnAlaSerProLeuIleSerProLeuAAspAAspGln	520
DB	1549	AACCTGAGCCCGACACCAAGAGGCTTCTCTTGATAGCCGCTGTAATGACCAAG	1608
QY	521	AlaCysProArgThrAAspAAspGluAAspGlyArgArgLysArgPheProThrAAspLys	540
DB	1609	GCCTGCCCGGACGAGATGAGAGTGAAGGCGCGAGAGAAAGATTTCCCACTGATTA	1668
QY	541	AlaTyrPheIleAlaLysGluValSerThrThrGlnAAspThrTyrLeuLysAAspLeuGlu	560
DB	1669	GGGTACTTATGCTTAAGAGATGCTTACACCGAGCGAACTATCTGAAAGATTCGGA	1728
QY	561	ValIleThrSerThrPheGlnSerThrValSerLysGluAAspAlaAAspProGluAlaLeu	580
DB	1729	GTTATCATCTTCGTGTTTCAAGACACATGAGCAAAAGAGAGCCCATGCCGAGAGACTG	1788
QY	581	LysSerLeuIlePheProAAspPheGluProLeuHsLysPheHsIleThrAAspPheLys	600
DB	1789	AAAAGTCTCATATTCGCCGAATTTTGAACCTTTGCAAAATTTCACTAATTTTTCAG	1848
QY	601	GluIleGlnGlnAAspLeuAAspLeuTrrGlyArgSerAAsnAlaGlnIleArgAAspTyr	620
DB	1849	GAAATTTGAGCAACGCTTCCCTGGGAAAGCCGCTCAAAAGCCCAATTCAGAGTTTAC	1908
QY	621	GlnArgIleGlyAAspValMetLeuLysAAsnIleGlnGlyMetLysHsIleLeuAAsnHs	640
DB	1909	CAAAGATGCGGAGATGCTAGCTGAAGACATTCAGGAGCATAGAGCACTGGCGGCTCAC	1968
QY	641	LeuTrrLysHsSerGlnAAspLeuAAspLeuAAspGlnGlyLysSerSerAAspArg	660
DB	1969	CTGTGGAACACACAGAGGCTTTGAGGCGCTTGGAATAGGATCAAGAGCTTCCCGCGG	2028
QY	661	LeuGluAAsnPheCysArgAAspPheGlnLeuGlnLysValCysTyrLeuProLeuAAsnThr	680
DB	2029	CTGGAGAACTTCTGCAGACATTTGAGCTGCAGAAAGTGTGTTACCTTACCGCTCAACACC	2088
QY	681	PheLeuLeuAAspProLeuHsIleArgLeuMetHsIleTyrLysGlnValLeuGlnAAspCys	700
DB	2089	TTCTCTCGGCGCACTGACCGGCTCATGCACTTACAGCAAGGTCTTGAAGCGGCTGTGC	2148

QY 701 LysHisProProSerHisAlaAspPheArgAspCysArgAlaAlaLeuAlaGluIle 720
 DB 2149 AAcACACACCCGCGAGCCAGCCGACTTCAGGAGCTGGCCAGCGGCTTGGCAGAAATC 2208
 QY 721 ThrGluMetValAlaGluLeuHisGlyThrMetIleIysMetGluAsnPheGlnIleLeu 740
 DB 2209 ACGGAGATGGGTGGACAGCTCCAGGTACGATGATCAAGATGGAATTTCCAAACCTG 2268
 QY 741 HisGluLeuIleValAspLeuIleGlyIleAspAsnLeuValAlaProGlyValGluPhe 760
 DB 2269 CAGGACTCAAGAAAGATTGATTGACATTCATCTTGCTCCGGGGAAGGAGATTC 2328
 QY 761 LlaArgGluGlySerLeuSerIleValGlyGluGluGlnIleAsnMetPhePhe 780
 DB 2329 ATCCGCTGGGACCTTCAGCAGCTCTCGGGGAAGGGGCTCCAGCAGCGCATGCTTCTC 2388
 QY 781 LeuPheAsnAspValLeuLeuIleThrSerArgIleuThrAlaSerAsnGlnPheIys 800
 DB 2389 CTGTTCAACGACGCTCGCTATACAGAGCGGGGGCTGACGGCTCCAAATCAGTTTAA 2448
 QY 801 ValHisGlyGluLeuProLeuIleGlyMetThrIleGluGluSerGluAspGluTyrGly 820
 DB 2449 GTCCACGCGGAGCTCCGCTCTATAGCATGACATTGAGAGAGCGAAGACGAGTGGGG 2508
 QY 821 ValProHisCysLeuThrLeuArgGlyGlnArgGlnSerIleIleValAlaAlaSerSer 840
 DB 2509 GTGCGCCACGCTGACCTCGGGGGCCAGGGGCGATCATCGTGGCCGCAAGTTCT 2568
 QY 841 ArgSerGluMetGluIleTyrValGluAspIleGluMetAlaIleAspLeuAlaGluIys 860
 DB 2569 CGGTCGAGATGAGAAATGGGTGAGAGACATCCAGATGGCCATTGACCTGGCGAAG 2628
 QY 861 SerSerSerProAlaProGluPheLeuAlaSerSerProProAspAsnIleSerProAsp 880
 DB 2629 AGCAGCGCGCCCGCTGAGTTCTGCGCCAGCGCCCTGCAACAAAGTCCCTGAT 2688
 QY 881 GluAlaThrAlaAlaAspGlnIleSerGluAspAspLeuSerAlaSerArgThrSerLeu 900
 DB 2689 GAAGGCCACCGGGGTGACAGAGATCAAGAGATACCTGAGCGCTCGCGACATCCCTG 2748
 QY 901 GluArgGlnAlaProHisArgGlyAsnThrMetValHisValCysTrpHisArgAsnThr 920
 DB 2749 GAGCGGACGCGCCGCGACCGCGCAACATGGTGCACGTGCTGCGACCGCAACACC 2808
 QY 921 SerValSerMetValaAspSerIleAlaValGluAsnGluLeuSerGlyAsnLeuLeu 940
 DB 2809 AGCGTCTCCATGGGACTTCAGCATCGCAGTGGAGAAATCAGTTGCTCGAAACCTGCTG 2868
 QY 941 ArgValPheIleValAsnSerAsnGlyTrpGlnIleLeuIleValAlaPheThrAsnPheCys 960
 DB 2869 AGAATAATTCAAACACACACCGGGTGGCAGAAAGTGGGGGTGCTGCACAAATTTTGC 2928
 QY 961 LeuPhePheIleIleValSerHisGlnAspAsnHisProLeuAlaSerLeuProLeuLeuGly 980
 DB 2929 CTGTTCTTCAAAATACACACGAGACATCATCCCTTGGCAGCCCTGCTGCTGGCC 2988
 QY 981 TyrSerLeuThrIleProSerGluSerGluAsnIleGlnIleValAspTyrValPheIleLeu 1000
 DB 2989 TACTCGCTCAACATCCCTCTGAGTCCGAGAAACATCCAGAAAGACTGCTTCAACCTG 3048
 QY 1001 HisPheIleSerHisValIleTyrIlePheArgAlaGluSerGluIleIlePheGluArgTyr 1020
 DB 3049 CACTTCAGAGTCCACGCTACTACTACAGGGCGGAAAGCAGATGACAGTTGAAAGGTGG 3108
 QY 1021 MetGluValIleArgSerAlaThrSerSerAlaSerArgProHisValIleSerHisIle 1040
 DB 3109 ATGGAAGTGAATCCGACATGGCCACAGCTCTGCTCGGAGCCCACTGTTGAGCCACAAA 3168
 QY 1041 GluSerLeuValIleTyr 1045
 DB 3169 GAGTCTCTGTGTAT 3183

RESULT 2
 AB008430
 LOCUS Homo sapiens mRNA for CDEP, complete cds.
 DEFINITION
 AB008430
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
 1 (sites)
 Koyano, Y., Kawamoto, T., Shen, M., Yan, W., Noshiro, M., Fujii, K. and Kato, Y.
 Molecular cloning and characterization of CDEP, a novel human protein containing the ezrin-like domain of the band 4.1 superfamily and the Dbl homology domain of Rho guanine nucleotide exchange factors
 Biochem. Biophys. Res. Commun. 241 (2), 369-375 (1997)
 9425278
 JOURNAL
 PUBMED
 REFERENCE
 AUTHORS
 Koyano, Y., Kawamoto, T. and Kato, Y.
 Direct Submission
 Submitted (22-OCT-1997) Takeshi Kawamoto, Hiroshima University School of Dentistry, Department of Biochemistry, 1-2-3 Kasumi Minami-ku, Hiroshima, Hiroshima 734, Japan
 (E-mail: k Kawamoto@ipc.hiroshima-u.ac.jp, Tel: 082-257-5688, Fax: 082-257-5629)
 FEATURES
 source
 1..3442
 /location/Qualifiers
 1..3442
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /cell_type="Chondrocyte"
 /tissue_type="cartilage"
 /dev_stage="embryo"
 49..3186
 /function="Rho Guanine Nucleotide Exchange Factor"
 /note="band 4.1 superfamily"
 /codon_start=1
 /product="CDEP"
 /protein_id="BA24267.1"
 /db_xref="GI:2766165"
 /translation="MSERIORPTPGSRLLAPENSGISTLERGQKPEPTSGKIVSIKI QMLDQVEAFVFORAPGVLDVACNHLNVEGDFGLSEFPDHKKITWLDLKPVI KOIRPKHVAVKVFVFPEDDTLOBELIRYLFALQVQDLQGLTCDNSTALILI SHIVQELIGDFEALDRHLAKNKYIPOODALEDKIVFEHNNIGQTPASDQLLEI ARLEWYGIIRLHPAKOREGKINLAVANTGILVFOGFTKINAFNMKVRLLSPKRRF LTKLRDANSAYODTLEPLMASRDPCKSPFKICVEHHAAPRLFEEDKPKPKPVLPARG SSFRSGRTQKQVLDYKSGGKIKVOPERKHKISIRLSAOPTELNSVLEDSQOS TSLTFEGBASPEGQSGCRGKEPKVAGSGSPSPAPRSPKGNQADDAASAAPTEE BEEVVNDRTQOSKPQPPQSTGSLTGSPIHSELVNSQGVAPANTVLSNLSPDQK ASPLIPLLDQACPTDDEDEGRARRFPDVKYVFLAKEVSTTERVYLKLEVITSWF OSTVSKEDKPEALKSLIPNPEPLKHTNPLFEKEIQRALMEGSMQIIRYORIG DWLKNIOGMKPEALHLMKHSRLELNGICYSRLELNCQPELOKQCVLPLNFTL LRPILRLMTRYKQVLERLCKHNPSSHPDPCRAALAEITBMVAQLGTMKMNPOUL HEIKKOLIGIDNLVVGREFIRLGSLSKSGKALQQRMPFLPNDVLLYTRGLTASL FKHGQILPLYGMTIIBESBDEWGVHCLITLGRQSIITVAASSEMEKWEIDQMAID LAEKSSPAPEPLASPPPNKSPDEATAADQESBEDLSASRTLELQAPRGTQMWY CHMRNTSVSVQPSIIVENOLSGNLKRFKNSGMQKLMVVFNFCLPFYKSHQDHP LSLPLIGVSTIRPSESNIOKQVYVKLHFKHSVYFRASEVTFRMMEVIRSATSS ASRPVLSKESLVY"
 3442
 /note="50 a nucleotides"
 polyA_site
 ORIGIN
 Alignment Scores:
 Pred. No.: 0
 Score: 5463.00
 Percent Similarity: 100.0%
 Best Local Similarity: 100.0%
 Query Match: 100.0%
 Length: 3442
 Matches: 1045
 Conservative: 0
 Mismatches: 0
 Indels: 0

DB: 5 Gaps: 0
US-09-555-342B-2 (1-1045) x AB008430 (1-3442)
QY 1 MetGlyGluIleGluGlnArgProThrProGlySerArgLeuGluValAProGluAAsnSer 20
DB 49 ATGGGAGAAATAGACGAGAGGCGACCCCGAGATCAAGCTGGGGGCCCCGGGAAATTCG 108
QY 21 GlyTLeSerThrLeuGluValArgGlyGluValAspProProThrProSerGlyValLeuVal 40
DB 109 GGGATACGATCTTGGAACGTGGACAGAGCCGCCCCCAACACCTTCAGGAAATCTCGTG 168
QY 41 SerTLeuLeIleGlnMetLeuAspPheThrGlnGluValAAspGluValAProGlnArgAla 60
DB 169 TCATCAAAATCCAGATGCTGATGACACCCAGAGAGCATTTGAAGTCCACAAAGAGCT 228
QY 61 ProGlyValValLeuLeuAspAlaValCyAsnHisLeuAsnLeuValGluGlyAspTyr 80
DB 229 CCGGGAGAGGTGCTGCTGATGACATGGTCMAACCTCAACCTCGTGGAAAGGTGACTAT 288
QY 81 PheGlyLeuGluPhePheProAspHisValValLeThrValTrPheLeuAspLeuValPro 100
DB 289 TTGGCTCGAGTTCTCTGATCAAAAGATCAAGGTGGCTGGATCTCTAAACCC 348
QY 101 LLeValLeuGlnLeuArgArgProLysHisLeuValValLysPheValValLysPhePhe 120
DB 349 ATTGTAAACAGATTGAAGGCCAAAGCACGTTGTTTAAGTTTGTGTGTAATTCCTT 408
QY 121 ProProAspHisThrGlnLeuGlnGluValLeuThrArgTyrLeuPheAlaLeuGlnVal 140
DB 409 CCGCTCGACACACACAACTCCAAAGAAAGACTCAAGAGTACTGTTGGCGCTGACAGTG 468
QY 141 LysGluAspLeuAlaGlnGlyArgLeuThrCyAsnAspPheSerAlaAlaLeuLeuIle 160
DB 469 AAGCAGAGACTTGGCTCAAGCAGAGTTGACGTGAAGACACAGCGCAGCTCTTGATT 528
QY 161 SerHisIleValGlnSerGluIleGlyAspPheAspGluValaLeuAspArgGluHisLeu 180
DB 529 TCACATGTTGCAATCTGAGATTGGGGATTTTGATGAAGCTTTGACAGAGACCTTA 568
QY 181 AlaLysAsnLysTyrTleProGlnGlnAspAlaLeuGluAspLysIleValGluPheHis 200
DB 589 GCAGAAAATTAATATACACTCTGACAAAGCAGCACTGAGAGCAAAATTCGGGAATTCAC 648
QY 201 HisAsnHisIleGlyGlnThrProAlaGluSerAspPheGlnLeuLeuGluIleAlaAsx 220
DB 649 CATTAACCACTTGGACAAACACAGCAGATCAAGATTTCCAGCTCTTAAGATTGCCCT 708
QY 221 ArgLeuGluMetTyrGlyIleArgLeuHisProAlaLysAspArgGluGlyThrLysIle 240
DB 709 CGGCTAGAGATGTATGGAATCGGTTGCACCGGSCCAAGSACAGGAGGACAGAGATC 768
QY 241 AsnLeuAlaValAlaAsnThrGlyIleLeuValPheGlnGlyPheThrLysIleAsnAla 260
DB 769 AATCTGGCCGTTGCCAAACAGGGAATTCATGTCTTCAAGGTTTCACTAAGATCAATGCC 828
QY 261 PheAsnTrpAlaLysValaArgLysLeuSerPheLysArgTyrAspPheLeuIleLysLeu 280
DB 829 TTCAACTGGGCCAAGGTGGAGAGCTGAGCTTCAAGAGAAAGGCTTTCTCAACAGCTC 888
QY 281 ArgProAspAlaAsnSerAlaTyrGlnAspThrLeuGluPheLeuMetAlaSerArgAsp 300
DB 889 CGGCGAGATGCCAATAGTGGTACAGGATACCTTGAATTCCTGATGGCCAGTCCGGAT 948
QY 301 PheCysLysSerPheThrLysIleCysValGluHisHisAlaPhePheAlaGluPheGlu 320
DB 949 TTCTGCAAGCTCTTCGAAATCTGTGTGAACATCATGCTCTTTAACACTTTTGAA 1008
QY 321 GluProLysProLysProLysProValLeuPheSerArgLysSerPheArgPheSer 340
DB 1009 GAGCCCAACCAAGCCCAAGCCGCTCTTTAAGCCGGGGGATCATATTCGGTCACT 1068
QY 341 GlyArgThrGlnLysGlnValLeuAspTyrValLysGluGlyGluHisLysValGln 360

DB 1069 GGTCCGAGCTCAGAGACAGGTTCTCAGATGTAAAGAGAGACATAGAAGGTGCGAG 1128
QY 361 PheGluArgLysHisSerTyrIleHisSerTleArgSerLeuAlaSerGlnProThrGlu 380
DB 1129 TTGGAAGGAAACACAGCAAGATTCATTTATCCGAGCCTTGCTTACAGCTTACAGAA 1188
QY 381 LeuAsnSerGluValLeuGluGlnSerGlnGlnSerThrSerLeuThrPheGlyGluGly 400
DB 1189 CTGAATTCGGAAGTGTGGAGCAGTCTCAGCAGAGACCAAGCTTATCTTGGAGAAAGGT 1248
QY 401 AlaGluSerProGlyGlyGlnSerCyAspArgGlyLysGluProLysValSerAlaGly 420
DB 1249 GCCGAATCTCCAGGGGGCCAGAGCTGCCGCCAGGAAAGAAACCGAAGGTTTCCGGGG 1308
QY 421 GluProGlySerHisProSerProAlaProArgArgSerProAlaGlyAsnLysGlnAla 440
DB 1309 GAGCCGGGGTCCGACCCGAGCCTTGCGCGAGAGAGAGCCCGGGTAAACAGCAGGCG 1368
QY 441 AspGlyAlaAlaSerAlaProThrArgGluGluGluValValLysAspArgThrGln 460
DB 1369 GAGGAGCGGCTCGGCGCCACGAGAGAGAGAGAGAGTGTTAAGATAGGACCCAG 1428
QY 461 GlnSerLysProGlnProProGlnProSerThrGlySerLeuThrGlySerProHisLeu 480
DB 1429 CAGAGTAAACTCAGCCCGCAGCCACAGCACAGGCTCTCACTGGCAGTCTCACTT 1488
QY 481 SerGluLeuSerValaAsnSerGlnGlyValAlaProAlaAsnValThrLeuSerPro 500
DB 1489 TCCGAGCTGTGTGAATCTGCAAGGGGAGAGGGCCCTGCGCAACGTGACTGTCTCC 1548
QY 501 AsnLeuSerProAspPheThrLysGlnAlaSerProLeuIleSerProLeuLeuAsnArgGln 520
DB 1549 AAGCTGAGCCCGGACACCAAGCAGAGGCTCTCCCTTGATCAGCCGCTGTGAATGACAG 1608
QY 521 AlaCyProArgThrAspAspGluAspGluGlyArgArgLysArgPheProThrAspLys 540
DB 1609 GCCTGCCCCGAGCAGCATGAGTGAAGGGCCGAGAGAAAGATTCACACTGTAATA 1668
QY 541 AlaTyrPheIleAlaLysGluValaSerThrThrGlnArgTyrTyrLeuLysAspLeuGlu 560
DB 1669 GGTACTTATAGCTAAGAGAGTGTCTACACCGAGCAATATTTGAAGATCTCGAA 1728
QY 561 ValIleThrSerTyrPheGlnSerThrValSerLysGluAspAlaMetProGluAlaLeu 580
DB 1729 GTTATCACTTCGTGTTTCAAGACAGTGAAGAAAGAGAGCCATGCGGAAAGCAGTG 1788
QY 581 LysSerLeuIlePhePheProAsnPheGluProLeuHisLysPheHisThrAsnPheLeu 600
DB 1789 AAAAGCTCATATTCGGAATTTTGAACCTTGGCAAAATTCATATCTAATTTTCTCAAG 1848
QY 601 GlnIleGluGlnArgLeuAlaLeuTrpGluGlyArgSerAsnAlaGlnIleArgAspTyr 620
DB 1849 GAATATTGACAAAGACTTCCCTGTGGAAAGCCGCTCAAAATGCCAAATCAGAGATTAC 1908
QY 621 GlnArgIleGlyAspValaMetLeuLysAsnIleGlnGlyMetLysHisLeuAlaAlaHis 640
DB 1909 CAAAGATCGGGAGATGTATGTGAAGAACTTCAAGGAGCATGAAGCACTGGGGCTCAC 1968
QY 641 LeuTrpLysHisSerGluAlaLeuGluValaLeuGluAsnGlyTleLysSerSerArgArg 660
DB 1969 CTGTGGAAACACAGCAGAGCCTTGGAGGCCCTGGAAATGGAAATCAAGAGCTCCGGGG 2028
QY 661 LeuGluAsnPheCyAspArgPheGluGluGlnLysValaCyAspTyrLeuProLeuAsnThr 680
DB 2029 CTGGAGAACTTCTGCAAGACTTTTGAAGCTGCAAGAGGTGTATTAACCGCTCAACACC 2088
QY 681 PheLeuLeuArgProLeuHisArgLeuMetHisTyrLysGlnValLeuGluArgLeuCy 700
DB 2089 TTCTCTCTGCGGCACTGACCGGCTCATGCACTAACAAGAGTCTTGAGCGGCTGTGC 2148
QY 701 LysHisHisProProSerHisAlaAspPheArgAspCyAspArgAlaAlaLeuAlaGluIle 720

Db	2149	AAAGCAACACCGCGCGAGCCAGCCGACCTTCAGAGGATCTCCGAGCCGGCTTTGGCAGAGATC	2208
Oy	721	ThrgIuMetValAlaGlnLeuHieglYThrMetIleYMetGluAsnPhenGlnYLeu	740
Db	2209	ACGGAGATGGTGGCAGCTCCAGCTCCAGATGATCAAGATGGAGAAATTTCCAGAAAGCTG	2268
Oy	741	HisGluLeuLysValAspLeuIleGlyIleAspAsnLeuValProGluYArgGluIuHe	760
Db	2269	CACGAAATCTCAAGAAAGATTTGATTTGGCAGATGACAAATCTTGTGGTTCTCGGAGAGGAGTTTC	2328
Oy	761	IleatrgLeuGlySerLeuSerYsLeuSerGlyLysGlyLeuGlnGluArgMetPheHe	780
Oy	2339	ATCCGTCCTGGGCAAGCTTCAGCAAGCTTCGGGGAAAGGGGCTTCAGCAAGCATGTTCTTC	2388
Db	781	LeuPheAsnAspValIleuLeuYrThrSerArgGlyLeuThrAlaSerAsnGlnPheLys	800
Db	2389	CTGTTCAACGAGAGCTCTGCTGATATACAGAGCCGGGGGCTGACGGCTCCAAATCAGTTTAA	2448
Oy	801	ValHisGlyGlnLeuProLeuYrGlyMetThrIleGluGluSerGluAspGluTrrpGly	820
Db	2449	GTCACCGGAGCAAGCTCCGCTCATGAGCATGACGATTTGAGAGAGCGAAAGACGAGGGGG	2508
Oy	821	ValProHisGlyLeuThrLeuAspArgGlyGlnArgGlnSerIleIleValAlaIaSerSer	840
Db	2509	GTGCCCACTGCTTACCTCCGGGGGCGAGCGGCGATCATGATGTGGCGCCAGTTCT	2568
Oy	841	ArgSerGluMetGluLysEtrpValGluAspIleGlnMetAlaIeAspLeuAlaGluLys	860
Db	2569	CGGTCCGAGATGGAGAGTGGGTTGAGGACATCCAGATGGCCATTGACTGGCGGAGAAAG	2628
Oy	861	SerSerSerProAlaProGluPheLeuAlaSerSerProProAsnLysSerProAsp	880
Db	2629	AGCAGCAACCCCCCGCCCTGAGTTCCTGGCGAGCAACCCCCCTGCACAAAGTCCCTGAT	2688
Oy	881	GluIaIeThrAlaIaAspGlnGluSerGluAspAspLeuSerAlaSerArgThrSerLeu	900
Db	2689	GAGGCAACCGGAGCTGACCAAGAGATCAAGAGATGACCTGAGGCGCTCCGGCACAATCGCTG	2748
Oy	901	GluArgGlnAlaProHisArgGlyYasnThrMetValHisValCysTrpHisArgYasnThr	920
Db	2749	GAGCGCCAGGCCCCCGCACCGCGGCAACAAATGATGATGCACTGTGCTGGCACCGCAACACC	2808
Oy	921	SerValSerMetValAspPheSerIleAlaValGluAsnGlnLeuSerGlyYasnLeu	940
Db	2809	AGCGCTCCATGTTGACTTCAGCATCCGACAGTGGAGATTCAGTTGTCTGAAACCTGCTG	2868
Oy	941	ArgLysPheLysAsnSerAsnGlyTrrpGlnLysLeuTrrpValPheThrAsnPheCys	960
Db	2869	AGGAATTCCAAAAACAGACCGGAGTGGAGAAAGCTGTGGTGGTTTACAAACTTCCTGC	2928
Oy	961	LeuPhePheYrLysSerHisGlnAspAsnHisProLeuAlaSerLeuProLeuGlnY	980
Db	2929	CTGTTCTTCTCAAAATCAACACGAGCAATCATCCCTTTCGACGCTCCTCTGCTCGGC	2988
Oy	981	TyrSerLeuThrIleProSerGluSerGluAsnIleGlnLysAspYrValPheLysLeu	1000
Db	2989	TACTGCTCACCATCCCTCTGAGTCCGAGAACATCCAGAAAGACTAGTGTTCAGAGCTG	3048
Oy	1001	HisPheLysSerHisValYrTrrpPheArgAlaGluSerGluYrThrPheGluArgTrrp	1020
Db	3049	CAC TTCAAAGTCCCAAGTCTACTACTCAGGGGGGAAACGAGTACACCTTGAAAGGGTG	3108
Oy	1021	MetGluValIleArgSerAlaThrSerSerAlaSerArgProHisValLeuSerHisLys	1040
Db	3109	ATGGAAGTGAATCCGAGTGGCCACCAAGCTCTGCTGCGAACCACGTTTAAGCCACAA	3168
Oy	1041	GluSerLeuValYr 1045	
Db	3169	GAGTCTCTGTGTAT 3183	
RESULT 3			
BC041595		3546 bp	mRNA linear PRI 08-MAR-2005
LOCUS			

DEFINITION	Homo sapiens FERM, RhocEF (ABHGEF) and placketerin domain protein 1 (Chondrocyte-derived), transcript variant 1, mRNA (CDNA clone MGC:54223 IMAGE:5443230), complete cds.
ACCESSION	BC041595
VERSION	BC041595.1
KEYWORDS	GI:27370806
SOURCE	MGC.
ORGANISM	Homo sapiens (human)
REFERENCE	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 3546)
AUTHORS	Strauberg RU, Reingold EA, Grouse LH, Derge JG, Klausner RD, Collins FS, Wagner J, Shermen CM, Schuller GD, Altschul SF, Zeeberg B, Buetow KH, Schaefer CF, Bhat NK, Hopkins RF, Jordan H, Moore T, Max SI, Wang J, Hsieh F, Diachenko L, Marusina K, Farmer JA, Rubin GM, Hong L, Stapleton M, Soares MB, Bonaldo MF, Casavant TL, Schreier TE, Brownstein WJ, Udell TB, Toshiyuki S, Carninci P, Prange C, Raha SS, Loquellano NA, Peters CJ, Abramson RD, Mullany SJ, Bosak SA, McEwan PJ, McKernan KJ, Malek JA, Gunnarone PH, Villalón S, Worley KC, Hale S, Garcia AM, Gay LJ, Hulyk SW, Richardson DK, Muzny DM, Sodergren EJ, Lu X, Gibbs RA, Fahey J, Helton E, Ketterman M, Madan A, Rodrigues S, Sanchez A, Whiting M, Madan A, Young AC, Shevchenko Y, Bouffard GG, Blakeley RW, Touchman JW, Green ED, Dickson MC, Rodriguez AC, Grimwood J, Schmutz J, Myers RM, Butlerfield YS, Krzyzanski MI, Skalska U, Smalins DE, Schnerch A, Schein JE, Jones SJ and Marra MA. Mammalian Gene Collection Program Team Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16699-16903 (2002)
CONSRPT	2 (bases 1 to 3546) Director MGC Project. Direct Submission
TITLE	Submitted (20-DEC-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
JOURNAL	NIH-MGC Project URL: http://imgc.ncl.nih.gov
PUBMED	Contact: MGC help desk
REFERENCES	Email: cgabbs-remail.nih.gov
AUTHORS	Tissue Procurement: ATCC
JOURNAL	CDNA Library Preparation: Rubin Laboratory CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN) DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland; Web site: http://www.nisc.nih.gov/ Contact: nisc.mgc@hgti.nih.gov Ahtker,N., Ayela,K., Beckertom-Sternberg,S.M., Benjamin,B., Blakeley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S., Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P., Hansen,N., Ho,S.-L., Karlins,B., Kwong,P., Latic,P., Legaebi,R., Maduro,Q.L., Mastello,C., Maestri,B., Mastrian,S.D., McCloskey,J.C., McComell,J., Pearson,R., Staaterjop,S., Thomas,P.J., Touchman,J.W., Tsirngson,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L., Young,A., Zhang,L.-H. and Green,E.D.
REMARK	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: http://image.lnl.gov Series: IRAL Plate: 44 Row: 1 Column: 24. Location/Qualifiers 1. 3546 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="MGC:54223 IMAGE:5443230" /tissue_type="Brain, astrocytoma, grade IV" /clone_id="NIH MGC_98" /lab_host="DH10B" /note="Vector: pOTB7" 1. 3546
FEATURES	
SOURCE	
gene	

/gene="FARP1"
/note="synonyms: CDEP, MG697400"
/db_xref="GeneID:10160"
/db_xref="MIM:602654"
266..3403
/gene="FARP1"
/codon_start=1
/product="FERM, RhodGF, and plectestin domain protein 1, isoform 1"
/protein_id="AAH41595.1"
/db_xref="GI:49898405"
/db_xref="GeneID:10160"
/db_xref="MIM:602654"
/translation="MGSEIQRPTPGSRILGAPENSGISTLERGQKPPETSGKLVSIKI
OMLDDIOEAEPVOPARPGKVLDAVCNHLNLVEGDYGLFEPHKKITWMLDLKPLV
KQIRPKHVVVKFVVKPRPDHLOEELTRYLPAIOVKODLAOGSLTQNDTSALLI
SHIVQSEIGDFDALDREHLAKKTYIQODALDKIVFHHNTGOTPAESDQLEI
ARLEWYGRILHAKDREGTKINLAVANTGILVQGTKINAFNAKVRKLSFKRRP
LILRDPANSAYODTLEFLMASRDFCKSFMKI CVENHAPFLRPEEPKPYLFSRG
SSEFSGRTQKQVLDYVEGGHKKVQERGSKHSIRSLASOPTLENSVLEQSOOS
TSLTPGSAESPQSGSCRGRKEPSPAPRSPAPRSPAGNQTQADSLAPTEP
EEVVDRTQOSKPOPPQESTGSLGSPHSELNVNSQGVAPANTTSPNLSPTDK
ASPLISPLNDQACPTDDEBGRKRFPYDKAVFLAKEVSTTERTYLKDLEVTSMF
QSTVSKEDAMPEALKSLIFENFEPLKFPHTNFKELIQRALMEGSMNADIRYQRIQ
DVALKNIQGMKGLAHLMKISEALELENGISKRSLKNEPCRDPELQKCYPLNFTL
LRLHLMKYKQVLERLCKHPPSHADPRCDALAEITMVAQLHGTMIKMNPKOL
HELKDLIGIDNLVPGREFRILGSLSKLSGKLAQRMPEFLFNDVLLYTSRGLTASNO
PKHKGQPLYGMTIEESBDEWKPHCTILRGQOSTIYVASSSEMEKVEDIQMAD
LAKSSPAPFLASSPDKNSPDEATPAQSEDDLSASRTLEKQAPKSHQNDH
CMHRTNSVSWPDSIAVENQLSGNLRFKFNNSGMQKLWVFTNFCILFPKXKQNDH
LASPLGLYSLTIPSESENTOKDYVFLHFKSHVYFRASEBETFERMEVIRSAVSS
ASRPVLSHRESLVY"

CDS

US-09-555-342B-2 (1-1045) x BC041595 (1-3546)
ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 3546
Score: 5463.00 Matches: 1045
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 5 Gaps: 0

US-09-555-342B-2 (1-1045) x BC041595 (1-3546)
QY 1 MeG1yGlu11leglugu1nArGProThrProG1ySerArGleuG1yAlaProG1yuaenSer 20
Db 266 ATGGAGAAATAGAGCAGAGCCGACCCACAGATCACGACTGGGGGCCCGGAAAAATTCG 325
QY 21 G1y11SerThrleuG1uArG1yG1nLyArProProBorThrProSerG1yLySleuVal 40
Db 326 GGGATCAGTACCTTGGAACTGGACAGAACCGGCCCAACCTTCAGAAAAATCTGTG 385
QY 41 Ser11elys11leglmetleuArpArpThrG1nG1uAlaPheG1yValProG1nArG1a 60
Db 386 TCCATCAAAATCCAGATGCTGATGACACCCAGAGAGCATTTGAAAGTTCCACAAAGGCT 445
QY 61 ProG1yLyArValleuLeuArpAlaValCyAsnH1sleuSnleuValG1uG1yArpTyx 80
Db 446 CCTGGGAAGGTGCTGCTGATGACGATTGGACCACTCACTCGTGGAGAGTACTAT 505
QY 81 PheG1yLeuG1uPheProArpH1sLyAr11eThArValTrPleuArpLeuLeuLyArPro 100
Db 506 TTTGGCTCGAGTTTCTGTATCAAAAAAGATCACGGTGTGGCTGATCTCTRAAAACC 565
QY 101 11eVal1yG1n11eArGArProLySh1eVal1yAlaLySphArVal1yLySphePhe 120
Db 566 ATTGTAACAGATTGAAGGCCAACAGCATTGTTGTTAAAGTTTGTGGAATTCCTTT 625
QY 121 ProProArpH1sThrG1nleuG1nG1uLeuThrArG1yThrPheArpHeAlaLeuG1nVal 140
Db 626 CCGCTGACCAACAACCTCCAAAGAACTCAAAAGTACCTGTTGGGCTGCGAGGTG 685
QY 141 LyG1nArpLeuAlaG1nG1yArGLeuThrCyAsnArpThrSerAlaAlaLeuLeu11e 160

Db 686 AAGCAGACTTGGCTCAAGCAGAGTGAAGTGAATGACACACGCGCAGCTCTTGATTT 745
QY 161 SerH1s11eValG1nSerG1u11leg1yArpPheArpG1uAlaLeuArpArG1uH1sLeu 180
Db 746 TCACACATTGTGCAATCTAGATTTGGGATTTTGAATGAAGCTTTGACAGAGCACTTA 805
QY 181 AlAlaYAsn1yAr11ePProG1nG1nArpAlaLeuG1uArpLyS11eValG1uPheH1s 200
Db 806 GCAAAAAATPAATACACTTACTGACGAAGCGACTAGAGACAAAATCGTGAATTTTAC 865
QY 201 H1sAsnH1s11leg1yG1nThrProAlaG1uSerArpPheG1nleuLeuG1u11eAlArG 220
Db 866 CATAACCAATTGMAAACAACAGCAGATCATGATTTCCAGCTCTTAGAGATTCGCCGT 925
QY 221 ArGLeuG1uMeCyArG1y11eArGLeuH1sProAlaLyArpArG1uG1yThrLyS11e 240
Db 926 CGGCTAGAGATGATGAATCCGGTTGCAACCGGCAAGGACAGGAAGCGCAAGATC 985
QY 241 AsnLeuAlaValAlaAsnThrG1y11eLeuValPheG1nG1yPheThrLyS11eAsnAla 260
Db 986 AATCTGCGCTGCCAACAACGGGAATTCATGTTTCAAGGTTTCAATGAATCAATGCC 1045
QY 261 PheAsnTrpAlaLyArValArG1yLeuSerPheLyArG1yArG1yPheLeu11eLySleu 280
Db 1046 TTCACTGGGCCCAAGGTGCGAAAGCTGAGACTTCAAGAGAAAGCGCTTTCATCAAGCTC 1105
QY 281 ArGProArpAlaAsnSerAlaTyrg1nArpThrLeuG1uPheLeuMeAlaSerArpArp 300
Db 1106 CGGCGAGATGCCAATGTGCTGACGAGATACCTTGAAATTCCTGATGGCAGATCGGGAT 1165
QY 301 PheCyLeSerPheThrPlyS11eCyArValG1uH1sAlaPhePheArGLeuPheG1u 320
Db 1166 TTCTGCAAGTCTTTCGAAAAATCTGTGTGAACATCATAGCTCTTTAGACTTTTGA 1225
QY 321 GluProLyArProLyArProLyArProValLeuPheSerArG1ySerSerPheArG1yPheSer 340
Db 1226 GAGCCCAACCAAAAGCCCAAGCCCGTCTTAAAGCGGGGTCATCATTTCCGTTCAAGT 1285
QY 341 G1yArGThrG1nleuG1nValleuArpTyrg1nValLyG1nG1yG1yH1sLyS11eValG1n 360
Db 1286 GGTGGAAGCTCAGAAAGAGGATTTCTGACTATGTTAAAGAAAGAGAACTAAAGAGGTGAC 1345
QY 361 PheG1uArG1yH1sSerLyS11eH1sSer11eArGSer11eArG1ySerG1nProThrG1u 380
Db 1346 TTTGAAGAAGACACGAAAGATTCATTCACGAGCTTCTTCAACGCTTACAGAA 1405
QY 381 LeuArpSerG1uValleuG1uG1nSerG1nG1nSerThrSer11eThrPheG1yG1uG1y 400
Db 1406 CTGAATTCGAAAGTCTGAGAGCTCTCAGACAGACACAGCTTCAATTTGAGAAAGGT 1465
QY 401 AlaG1uSerProG1yG1yG1nSerCyArGArG1yLyG1uPProLyArValSerAlaG1y 420
Db 1466 GCCGAATTCACAGGGGGCCAGAGCTGCGGCGAGGAAGAAACCAAGGTTTCCGCCGGG 1525
QY 421 GluProG1ySerH1sArProSerProAlaProArGArGSerProAlaG1yAsnLyS1nAla 440
Db 1526 GAGCCGGGGTCCGACCCGACCTGCGCGAGGAAGAACCCCGGGGTAAACAAGCGGCG 1585
QY 441 AspG1yAlaAlaSerAlaProThrG1uG1uG1uG1uVal1yLySArpArGThrG1n 460
Db 1586 GACGAGACCGCTCGCGGCCCAACGAGAAAGAGAAAGTCTGTAAAGGTAGAACCAAG 1645
QY 461 G1nSerLyArProG1nProProG1nProSerThrG1ySerLeuThrG1ySerProH1sLeu 480
Db 1646 CAGAGTAACTTCAAGCCCCCGGACGCAACACAGCGCTCCCTGACTGACGTTCACTT 1705
QY 481 SerG1uLeuSerValAsnSerG1nG1yValAlaProAlaAsnValThrLeuSerPro 500
Db 1706 TCCGAGCTGTCTGTAACTTCGAGGGGGAGTGGCCCTTCGCAACGTGACTGTCTCCC 1765
QY 501 AsnLeuSerProArpThrLyS1nAlaSerProLeu11sSerProLeuLeuArpG1n 520

Dp	1766	AACTGACGACCCCGACACCAAGCAGGCTCTCCCTTGATGACCCGCTGCTGAATGACGAG	1823
Qy	521	AlAcYwPtoArqThrAspAspGluAspGluGlyAspArgPheProThrAspLys	540
Dp	1826	GCCTGCCCCCGGACGCGATGAGGATGAGGGCCGGAAGAAAGATTCCCAACTGATGAAA	1885
Qy	541	AlAtYrPheIleAlaLysGluValSerThrTrpGluGlyThrTyrLeuLysAspLeuGlu	560
Dp	1886	GCGRATCTTCATAGCTAAAGAAAGTGTCTAACACCGAGCAACAATATCTGAAGATCTCGAA	1945
Qy	561	ValIleThrSerTrpPheGlnSerThrValSerLysGluAspAlaMetProGluAlaLeu	580
Dp	1946	GTTATCACTTGTGTGTTTCAGAGCACATGAGCAAAAGAGACGCCATCCGGAAGCACTG	2005
Qy	581	LysSerLeuIlePheProAspPheGluProLeuHisLysPheHisThrAspPheLys	600
Dp	2006	AAAGGTCTCATATTCCCGAATTTTGAACTTTTGCACAAAATTTTCATTAATTTTCTCAAG	2065
Qy	601	GluIleGluGlnArgLeuAlaLeuTrpGluGlyArgSerAsnAlaGlnIleArgAspTyr	620
Dp	2066	GAATTTGAGCAACGACTTGCCCTGTGGGAAGCCGCTCAATGCCCAATGAGATTC	2125
Qy	621	GlnArgIleGlyAspAspValMetLeuLysAsnIleGlnGlyMetLysHisLysLeuAlaHis	640
Dp	2126	CAAAAGATCGCGGATGTATGCTGTAAGAACATTCAGGGCAATGAAGACCTTGCGGCTCAC	2185
Qy	641	LeuTrpLysHisSerGluAlaLeuGluValaLeuGluAsnGlyIleLysSerSerArgArg	660
Dp	2186	CTGTGGAAGACACAGAGAGGCGCTTGGAAGCCCTGGAGAAATGGAATCAAGAGCTCCCGGCGG	2245
Qy	661	LeuGluAsnPheCysValArgAspPheGluLeuGlnLysValCysTyrLeuProLeuAsnThr	680
Dp	2246	CTGGAAGAACTTCTGACAGAGACTTTGACGTGACGAAGGTGTCTTCACTCCGCTCAACACC	2305
Qy	681	PheLeuLeuAspProLeuHisArgLeuMetHisTyrTyrLysGlnValaLeuGluArgCys	700
Dp	2306	TTCCTCTCGCGGACATGACCGGCTACTAGCACTAACAGAGGCTCTGGAAGGCTGTGCG	2365
Qy	701	LysHisHisProProSerHisAlaAspPheArgAspCysArgAlaAlaLeuAlaGluIle	720
Dp	2366	AAACACACACCCCGGAGGCCACGCCGACTTCAGGAGACTGCCAGCGCTTTGGCAGAGATC	2425
Qy	721	ThrGluMetValaAlaGlnLeuHisGlyTyrMetIleLysMetGluAsnPheGlnLysLeu	740
Dp	2426	ACGGAAGATGTGGGCACAGACTCACAAGGTTCAGATGATCAAGATGGAGAAATTTCCAAAGCTG	2485
Qy	741	HisGluLeuLeuLysValAspLeuIleGlyLysLeuAsnLeuValProGluArgGluPhe	760
Dp	2486	CACCAACCTCAAGAAAGATTGATTTGACATTGACAAATCTTTGTGTCCGGGAAGGAGAGTTC	2545
Qy	761	IleArgLeuGlySerLeuSerLysLeuSerGlyLysGlyLeuGlnGlnArgMetPhePhe	780
Dp	2546	ATCCGTCGTGGGCAAGCTTCAGCAAGCTCTCGGGGAAGGGGCTTCAGCAGCGAGTGTCTTC	2605
Qy	781	LeuPheAsnAspValaLeuLeuTyrHisSerArgGlyLeuThrAlaSerAsnGlnPheLys	800
Dp	2606	CTGTTTCAACGACGTCCTGCTATACACGAGCGGGGGCTGACGCGCTCAATCAGTTTAAA	2665
Qy	801	ValHisGlyGlnLeuProLeuTyrGlyMetThrIleGluGluSerGluAspGluTrpGly	820
Dp	2666	GTCACACGGGACGCTCCCGCTTATGGCAATGACATGAGAGACCGAAGACGAATGGGGG	2725
Qy	821	ValProHisCysLeuThrLeuArgGlyValaArgGlnSerIleIleValAlaSerSer	840
Dp	2726	GTCGCCCACTGCTGACCTCCGGGGCCAGCGGCAGTCAATCATCTGTGGCCGCAATTTCT	2785
Qy	841	ArgSerGluMetGluLysTrpValGluAspIleGlnMetAlaIleAspLeuAlaGluLys	860
Dp	2786	CGGTCGAGAGATGGAGAAATGGGATTGAGACATCCAGATGGCCATTGACCTGGCGGAAGAG	2845
Qy	861	SerSerSerProAlaProGluPheLeuAlaSerSerProProAspAsnLysSerProAsp	880
Dp	2846	AGCAGCAACCCCGGCTGAGTCTCTGGCCAGCAGCCCCCTGCAACACMACTCCCTGAT	2905

QY		881	GluIaArGlnAlaAaAGNGlUSeRGlUbApSbPLeuSerAlAsErArgThSerLeu	900
Dd		2906	GAAgCACAAGCGGGCTGACCAAgAGTACAAGATACCTTGAGCGGCTCGGCACATCCGTG	29655
QY		901	GluArGlnAlaProHIsarGIyASnThrMeCVaHIsvAlCySTrPHIsarGAasnThr	920
Dd		2966	GAGGCCACAGGCCCGGCACCAGCATGATGGACCGTGTGTCGGCACCGCAACACC	30255
QY		921	SerValSerMetValAspPheSerILeAlaValGluangInLeuSerGIyASnLeuLeu	940
Dd		3026	AGCGCTCTCAAGCGGACTTCAGAGATCGCAATGGAGAATCAGTTGTCTGGAAACCTGCTG	30855
QY		941	ArgLySPheLyASanSerASenGIyTrGlnLyLeuTrPVaIVaIPhetrAsnPheCyS	960
Dd		3086	AGAAATTCAAATAACAGAACGGGTGGCAAGACTGTGGGTGTTCACAAACTTCTGC	31455
QY		961	LeupPhePheryLySPserHIsGLInAsPAsnHIsAProLeuaIasErLeuProLeuLeuGIy	980
Dd		3146	CTGTTCTTCTTAACAATACACACAGAGAAATCATCCCCCTGGCAGCTGCGCTGCGGC	32055
QY		981	TyrSerLeuThrILeprosERglUsErGIuAnIllegInLyAsPrTyValPheLyLeu	10000
Dd		3206	TACTCGCTCACCAATCCCCTCTGAGCTCGAGAACTACAGAAAGACTACGTGTTCAGACTG	32655
QY		1001	HIsPheLySPserHIsvalTYrTYrPheArgHIsagIUsErGIuTYrThrPheGIuArGYrP	10200
Dd		3266	CAC TTCANAGTCCACAGCTTACTTACTTCAGGGCGGAAAAGCGATACAGTTCCAAGGTGG	33255
QY		1021	MetGIuAlIIearGSerAlaThrSerSerAlaSerArgProHIsvalIleuSerHIsLyS	10400
Dd		3326	ATGGAAGTGAATCCGACAGGCCACAGCTCTGCGGACCCCAAGTGTGAGCCACAAA	33855
QY		1041	GIUsErLeuValTYr 1045	
Dd		3386	GAGTCTCTTGTTGTTAT 3400	
RESULT 4				
CS130723				
LOCUS	CS130723	5028 bp	DNA	linear PAT 02-AUG-2005
DEFINITION	Sequence 9 from Patent WO2005064009.			
ACCESSION	CS130723			
VERSION	CS130723..1 GI:71792793			
KEYWORDS				
SOURCE	.			
ORGANISM	Homo sapiens (human)			
REFERENCE	1			
AUTHORS	Obermoff, T.F.			
TITLE	Classification of cancer			
JOURNAL	Patent: WO 2005064009-A 9 14-JUL-2005; Atos Applied Biotechnology APS (DK) Location/Qualifiers			
FEATURES	source			
	1..5028			
	/organism="Homo sapiens"			
	/mol_type="unassigned DNA"			
	/db_xref="taxon:9606"			
	/note="FBSM, RhogEP (ARRGEF) and plectatrin domain pro"			
ORIGIN				
Alignment Scores:				
Pred. No.:	0	Length:	5028	
Score:	5463.00	Matches:	1045	
Percent Similarity:	100.0%	Conservative:	0	
Best Local Similarity:	100.0%	Mismatch:	0	
Query Match:	100.0%	Indels:	0	
DB:	2	Gaps:	0	
QY	US-09-555-342B-2 (1-1045) x CS130723 (1-5028)			
	1 MetGIyGlnIlegInArgProThrProGIySerArgLeuGIyAlaProGIuAnSer 20			

Db 337 ATGGAGAAATAGAGAGAGCGGACCCGAGATCAAGTGGGGGGCCCGAAAATTGG 396
Qy 21 GYIILSerThrLeuGlnuArgGlyGlnLysProProBorThrProSerGlyLysLeuVal 40
Db 397 GGATCATGTAACCTTGGAACGTGACAGAAACCGCCCAACACCTTCAGAGAAAATCTGTG 456
Qy 41 SerLysLysIleGlnMetLeuAspAspThrGlnGlnuLysPheGlnuValProGlnuArgLys 60
Db 457 TCATCAAAATCCAGATGCTGGATGACACCCAGAGAGCATTTGAAGTTCACAAAGAGCT 516
Qy 61 ProGlyLysValLeuLeuAspLysValCysAsnHisLeuAsnLeuValGlnLysAspTyr 80
Db 517 CTGGGAGAGGTGCTGGCTGATGATGACACCACTCAACCTCGTGGAGAGGTGACTAT 576
Qy 81 PheGlyLeuGlnuLysProAspHisLysLysValIleThrValTyrLeuAspLeuLysPro 100
Db 577 TTTGGCTCAGAGTTTCTGTATCAAAAGATCAAGGTGTGGCTGGATCTCTTAAACCC 636
Qy 101 ILValLysGlnIleAspArgProLysHisValValLysPheValValLysPhePhe 120
Db 637 ATTGTGAACAGATTAGAGGCGCAAGACGTTGTGTTAAGTTTGTGGTGAATTTCTTT 696
Qy 121 ProProAspHisThrGlnLeuGlnuGlnuGlnuLysThrArgTyrLeuPheLysLeuGlnuVal 140
Db 697 CCGCTGACCAACACAACTCCAAGAAAGACTCAAGAGTACCTGTTCGGCTGCAAGTGTG 756
Qy 141 LysGlnuAspLeuLysGlnuArgLysLeuThrCysAsnAspThrSerLysAlaLeuLeuLys 160
Db 757 AACAGAGACTGTGCTCAAGGACAGGTGACGTGTAAATACACAGCGAGCTCTTGATT 816
Qy 161 SerHisValValGlnSerGlnuIleGlyAspPheAspGlnuLysLeuAspArgGlnuHisLeu 180
Db 817 TCACACTTGTGCAATGTGATGATTTGGGGAATTTGATGAAGCTTGGACAGAGAGCACTTA 876
Qy 181 AlaLysAsnLysTyrIleProGlnuAspLysLeuAspLysLysValGlnPheHis 200
Db 877 GCAAAATAATATACATCACTCAGCAAGACGCACTAGAGAGCAAAATCGGGAATTTAC 936
Qy 201 HisAsnHisValIleGlnuLysProLysGlnuLysAspPheGlnuLeuGlnuLysAlaAsp 220
Db 937 CATTAACCACTTGTGACAAACACAGACAGATCAGATTTCCAGCTCTAGAGATTTGCCCT 996
Qy 221 ArgLeuGlnuMetTyrGlyLysArgLysLeuHisProAlaLysAspArgGlnuTyrLysLys 240
Db 997 CGGCTAGAGTGTATGGAAATCCGGTTGCACCCGGCCAGAGACAGAGGACAGAAAGATC 1056
Qy 241 AsnLeuAlaValAlaAsnThrGlyLysLeuValPheGlnuLysPheThrLysLysAsnAla 260
Db 1057 AATCTGGCCGTTGCCAAACAGGGAATCTAGTGTTCAGGGTTTCACTAAGATCAATGCC 1116
Qy 261 PheAsnThrAlaLysValArgLysLeuSerPheLysArgLysArgPheLeuLysLeu 280
Db 1117 TTCAACTGGGGCAAGGCGGAGCTGAGCTTCAAGGAGAAAGCGTTTTCATCAAGCTC 1176
Qy 281 ArgProAspLysAsnSerAlaTyrGlnuAspThrLeuGlnuLysPheLeuMetAlaSerArgAsp 300
Db 1177 CGGCCAGATGCCAATAGTGGCTACAGAGATCACTTGAATTTCTGATGGCAGATCCGGAT 1236
Qy 301 PheCysLysSerPheThrLysLysCysValGlnuHisAlaLysPheArgLeuPheGlnu 320
Db 1237 TTCTGCAAGCTCTTGGAAAATCTGTGTGAACATCATGCTCTTTAAGCTTTTGGAA 1296
Qy 321 GlnuProLysProLysProLysProValLeuPheSerArgLysSerPheArgPheSer 340
Db 1297 GAGCCCAACCAAGCCCAAGCCGCTCTTTAGCGGGGGTCAATCATTTGGTTCAAGT 1356
Qy 341 GlyArgThrGlnuLysGlnuValLeuAspTyrValLysGlnuLysLysValGlnu 360
Db 1357 GGTCCGAGCTCAGAAAGAGGTCTTCTGACTATGTAAAGAGAGACATTAAGAAAGGTGACG 1416
Qy 361 PheGlnuArgLysHisSerLysLysLeuSerIleArgSerLeuAlaSerGlnuProThrGlnu 380

Db 1417 TTGAAAGAGACACAGCAAGATTCAATCTATCCGAGGCTTGTCAAGCTTACAGAA 1476
Qy 381 LeuAsnSerGlnuValLeuGlnuLysGlnuLysThrSerLeuThrPheGlnuGlnuLys 400
Db 1477 CTGAATTCGAAAGTGTGGAGCAGCTTCAGCAGAGACAGGCTTATCTTGGAGAAAGT 1536
Qy 401 AlaGlnuSerProGlyLysLysLeuSerCysArgArgGlyLysGlnuProLysValSerLysLys 420
Db 1537 GCGGAATCTCCAGGGGGCCAGAGCTGCGCGGAGGAAAGAAACCAAGAGTTTCCGCGGG 1596
Qy 421 GlnuProGlySerHisProSerProAlaProArgArgSerProAlaLysLysGlnuAla 440
Db 1597 GAGCCGGGTGCCACCGGAGCCCTGTGCGAGAGAAAGCCCGCGGTAAACAAGAGGGG 1656
Qy 441 AspGlyAlaAlaSerLysProThrGlnuGlnuGlnuGlnuValValLysAspArgThrGlnu 460
Db 1657 GAGGAGCCGCTCCGCGCCCAAGAGAAAGAGAGAGTGTTAAGATAGAGACCCAG 1716
Qy 461 GlnuSerLysProGlnuProProGlnuProSerThrGlySerLeuThrGlySerProHisLeu 480
Db 1717 CAGAGTAACTCAGCCCGGAGCCCAAGACACAGGCTCTCTGACTGGCAGTCTTCACTT 1776
Qy 481 SerGlnuLysSerValAsnSerGlnuGlyValAlaProAlaAsnValThrLeuSerPro 500
Db 1777 TCGAGCTGTCTGTAACTCGAGAGGGGAGTGGCCCTGCGCAACGTGACTTGTCTCC 1836
Qy 501 AsnLeuSerProAspThrLysGlnuLysSerProLeuLysSerProLeuAsnAspGlnu 520
Db 1837 AACCTAGCCCGGACACCAAGAGGCTCTCTCTTATATAGCCCGCTGTGAATGACAG 1896
Qy 521 AlaCysProArgThrAspAspGlnuAspGlnuArgArgLysArgPheProThrAspLys 540
Db 1897 GCCTGCCCGGAGCAGATGAGATGAGGGCCGGAGAGAAAGATTTCCCACTGATAA 1956
Qy 541 AlaTyrPheIleAlaLysGlnuValSerThrThrGlnuArgThrTyrLeuLysAspLeuGlnu 560
Db 1957 GGCTACTTATAGCTTAAGAGAGTGTCTACCCAGGAGCAATCTGTGAAGATTCGAA 2016
Qy 561 ValIleThrSerThrPheGlnuSerThrValSerLysGlnuAspAlaMetProGlnuAlaLeu 580
Db 2017 GTTATACCTTGTGTGTTTCAAGACACAGTGAAGCAAGAGAGAGAGAGAGAGAGAGAG 2076
Qy 581 LysSerLeuLysPheProAspPheGlnuProLeuHisLysPheHisThrAsnPheLeuLys 600
Db 2077 AAAAGTCTATATCCCGAATTTTGAACCTTGGCAAAATTTCAATCTAATTTTCAAG 2136
Qy 601 GlnuIleGlnuLysArgLeuAlaLeuTyrGlnuArgSerAsnAlaGlnuLysArgAspTyr 620
Db 2137 GAAATTAAGCAACGACTTCCCTGTGGAGAGGCGGCTCAAAATGCCAAATCAGAGATTAC 2196
Qy 621 GlnuArgIleGlyAspValMetLeuLysAsnIleGlnuLysMetLysHisLeuAlaHis 640
Db 2197 CAAGAGATGGGAGATTCATGCTGAAGAACATTCAGGGGATGAAGACCTTGGGGCTCAC 2256
Qy 641 LeuTyrLysHisSerGlnuLysLeuGlnuAlaLeuGlnuAsnGlyLysLysSerSerArgArg 660
Db 2257 CTGTGGAACAACAGGAGGCTTGGAGGCTCGAGAAATGAAGATCAAGAGCTCCCGGG 2316
Qy 661 LeuGlnuAsnPheCysArgArgPheGlnuLysValCysTyrLeuProLeuAsnThr 680
Db 2317 CTGGAAACCTTCTGCAAGACTTGTGAGTCGAGAAAGTGTGTAACCTTACCGCTCAACACC 2376
Qy 681 PheLeuLeuArgProLeuHisArgLeuMetHisTyrLysGlnuValLeuGlnuArgLeuCys 700
Db 2377 TTCTCTTCGGGCACTGACCGGCTCATGCACTACAAAGAGTCTTGAAGCGGCTGTGC 2436
Qy 701 LysHisHisProProSerHisAlaAspPheArgAspCysArgAlaAlaLeuAlaGlnuLys 720
Db 2437 AAACACACACCGCGGAGCAGCCGACTTCAGAGGACTGCGGAGCCGCTTGGCAGAGATC 2496
Qy 721 ThrGlnuMetValAlaGlnuLeuHisGlyThrMetLysMetGlnuAsnPheGlnuLysLeu 740
Db 2497 ACGAGAGTGTGCAAGCTTCAAGGTGATGATCAAGATGAGAAATTTCCAGAAAGCTG 2556

Qy	741	HHGIHLEUPLYVAAPLEUHLIGLYLPHAPHEUVALVALPROGLYARGLUPHE	760
Dp	2557	CACGAACCTCAAGAAAGATTGATTTGGCATGACAATCTTGTTGGTCCGGAAAGGAGTTC	2616
Qy	761	ILearGLEUGLYSErLEuSErLYeUSeRGlyLYeGlyLEuGInGLARMePhePhe	780
Dp	2617	ATCCGTCTGGGCAAGCTCAGCAAGCTCTCGGGAAAGGGGCTCCAGACGCGATGTTCTTC	2676
Qy	781	LEuPheAaAPVALLEuLYrThrSErARGLYLeUTHrALASerAMGLPheLYs	800
Dp	2677	CTGTTCAAACGACGTCTCTGTATACACGAGCCGGGGGCTGACGGGCTCCAACTAGTTTAAA	2736
Qy	801	VALHHGIYGLNLEuPROLEuTYrGLYMeThrLIGLYuSErGLYARGLUTrGLY	820
Dp	2737	GTCCACGGGCAAGCTCCCGCTCTATGGCATGACGATTGAGAGAGGAAGACAGATGGGGG	2796
Qy	821	VALPRonHsCYbLEuThrLEuARGLYGLNARGLInSErLEIleVALAlASerSEr	840
Dp	2797	GTGCCCACTGCTGACCTCCGGGGCCAGCGGCGATCACTCTGTGGCCGACATGTTCT	2856
Qy	841	ARgSErGLuMErGLuLYeTrpVALGLuAPrLEGLMeTAlAlAeAPrLEuALGLuLYs	860
Dp	2857	CGGTCCGAGATGGAGAACTGGATTGAGACATCCAGATGGCCATGACTGGCGGAGAG	2916
Qy	861	SErSErSErPRonALPRoGLYpHeuALASerSErPRoPRoARhLYSeSErPRoAR	880
Dp	2917	AGCAGCACCGCCCGCCCTGAGTCTCTGGCCAGACGCCCCCTGCACMAAGTCCCTGAT	2976
Qy	881	GLuALAThrALALAAERGLNGLYuSErGLuAPrLEuSERALASerArThrSErLEu	900
Dp	2977	GAACCCACCGCGGCTGACCAAGAGTACAGATGACTGAGCGCTCGCGACATCGCTG	3036
Qy	901	GLuARGLNALPRonHsARGLYAAvThrMeTVALHsVALCYeTrpHsARGLAAvThr	920
Dp	3037	GAGCGCCAGGCCCCCGCACCGGCGACAACTAGTGCACGTGCTGGCACCGCAACCC	3096
Qy	921	SErVALSErMeTVALAPrPheSErLEIleVALGLuANGInLEuSErGLYANLEuLEu	940
Dp	3097	AGCGTCTCAAGTGTGACTCTGACGATCCGACGTGAGATCAAGTTGTCTGAAACCTGTGT	3156
Qy	941	ARGLYpHeLYAAvSErSErANGLYrGLNLYeUTrpVALALPheThrASpPheCYs	960
Dp	3157	AGGAAATTCAAAACACGACCGGGTGGCGAAGCTGTGGGTGTTTCAAACTTCTGC	3216
Qy	961	LEuPhePheTYrLYSeSErHsGLNAPrANhIAPrOLEuALASerLEuPROLEuENGLY	980
Dp	3217	CTGTTCCTCTCAAAATCACACAGGACAATATCCCTTGCAGGCTGCTCTGTGGCC	3276
Qy	981	TYrSErLEuThrLIErPROSErGLuSErGLuANhILEGLYbAPrTYrVALPheLYeLEu	1000
Dp	3277	TACTTCGCTCACATCTCCCTCTGAGTCCGAGAACTCCGAAAGACTAGTTCMAAGTGG	3336
Qy	1001	HIAPheLYSeSErHsVALTYrTYrPheARGALGLuSErGLYrThrPheGLuARTrp	1020
Dp	3337	CAC TTCAGTCCACGCTTACTTACTTCAAGGGCGAAACGAGTACACCTTGAAAGGTGG	3396
Qy	1021	MeTGLuVALIleArGSErALThrSErSErALASerARPRonHsVALLEuSErHsLYs	1040
Dp	3397	ATGGAAGTATCCGAGTGCACCAACGCTGCTGCGAACCCCAAGTGTGATGATCACAAA	3456
Qy	1041	GLuSErLEuVALTYr 1045	
Dp	3457	GAGTCTCTTGAT 3471	

RESULT	5				
LOCUS	BD231200				
DEFINITION	BD231200 Human cytoskeleton associated proteins.	4687 bp	DNA	linear	PAT 17-JUL-2003
ACCESSION	BD231200				
VERSION	BD231200.1	GI:33040970			
KEYWORDS	JP 200252676-A/2.				
SOURCE	Homo sapiens (human)				

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
ORGANISM	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo.			
	1 (bases 1 to 4687)			
	Lai, P., Tang, T. Y., Yue, H., Hillman, J. L., Bandman, O., Corley, N. C., Huenig, K. J., Patterson, C., Azimzai, Y. and Baughn, M. R.			
	Human cytoskeleton associated proteins			
	Patent: JP 200256076-A 2 20-AUG-2002;			
	INCYTE PHARMACEUTICALS INC			
	OS Homo sapiens (human)			
	PN JP 200256076-A/2			
	PD 20-AUG-2002			
	PF 17-SEP-1999 JP 2000574254			
	PR 18-SEP-1998 US 60/172266, 27-APR-1999 US 60/131321 PI			
	PRETTI LAL, TOM Y TANG, HENRY YUE, JENNIFER L HILLMAN, OLGA PI			
	BANDMAN,			
	PI NEIL C CORLEY, KARL J GUEGLER, CHANDRA PATTERSON, YALDA AZIMZAI,			
	PI MARIAM R BAUGHN			
	PC C12N15/09, A61K38/00, A61K45/00, A61P1/16, A61P3/00, A61P3/06 PC			
	, A61P3/10, A61P5/00,			
	PC A61P3/38, A61P7/00, A61P7/06, A61P9/00, A61P9/10, A61P11/06 PC			
	, A61P13/00, A61P13/12,			
	PC A61P17/06, A61P19/02, A61P19/10, A61P21/04, A61P25/00, A61P25/00,			
	PC A61P25/04,			
	PC A61P25/14, A61P25/16, A61P25/28, A61P29/00, A61P31/04, A61P31/12,			
	PC A61P31/18,			
	PC A61P33/00, A61P35/00, A61P35/02, A61P37/00, A61P43/00, C07K14/47,			
	PC C07K6/18,			
	PC C12N1/15, C12N1/19, C12N1/21, C12N5/10, C12P21/02, C12P22/08 PC			
	, C12Q1/68, G01N33/15,			
	PC G01N33/50, G01N33/53, G01N33/56, C12N5/00, C12N5/00,			
	PC A61K37/02			
	CC Incyte ID No: 1320252			
	CC Incyte			
	FM Key			
	FT source			
	1. .4687			
	/organism='Homo sapiens (human)'			
	/location/Qualifiers			
	1. .4687			
	/organism='Homo sapiens'			
	/mol_type='genomic DNA'			
	/db_xref='taxon:9606'			
ORIGIN				
Alignment Scores:				
Pred. No.:	0	length:	4687	
Score:	5459.00	Matches:	1044	
Percent Similarity:	99.9%	Conservative:	0	
Best Local Similarity:	99.9%	Mismatches:	1	
Query Match:	99.9%	Indels:	0	
DB:	2	Gaps:	0	
US-09-555-342B-2 (1-1045) x BD231200 (1-4687)				
QY	1 MetGlyGluIleGluGlnArgProThrProGlySerArgLeuGlyAlaProGluAsnSer 20			
Db	90 ATGGAGAAATGAGACAGAGCGCCGACCCAGATATCGACTGGGGGCCCCCGAAAATTCG 149			
QY	21 GlyIleSerThrLeuGlnArgGlyGlnLysProProPioThrProSerGlyLysLeuVal 40			
Db	150 GGGATCAGATACCTTGGAACGTGTGACAGAGCCGCCCCCAACCTTCAGAGAAATTCGTG 209			
QY	41 SerIleLysIleGluMetLeuAspAspThrGlnGluAlaPheGluValProGlnArgAla 60			
Db	210 TTCATCAAAATCCAAATGCTGGATGACACCGAGGCAATTGGAAATTCACAAAGAGCT 269			
QY	61 ProGlyLysValLeuLeuAspAlaValCysAsnHisLeuAsnLeuValGluGlyAspTyr 80			
Db	270 CCTGGAAAGAGTCTGCTGGATGCAAGTTTGCACCAACCTCAACTGCTGGAAGGTGACTAT 329			
QY	81 PheGlyLeuGluPheProAspHisIleLysLysIleThrValTTPLeuAspLeuLeuLysPro 100			
Db	330 TTGGCCCTTCGAGTTTCGATCACAAAGATCACGATGTGGCTGGATCTCCTTAAAGCC 389			

QY 101 ILeValIysGlnIleArgArpProLYsHLeValValIlyspHeValIlyspHePhe 120
Db 390 ATTGTGAACAGATTAGAGGCCAAGACAGTTGTTTAAAGTTTGAGTGAATTTCTTT 449
QY 121 ProProAspHLeThrGlnLeuGlnGluLeuThrArgTYrLeuPheAlaLeuGlnVal 140
Db 450 CCGCTGACCAACAATCCCAAGAAAGACTCAAGAGTACCTGTTCCGGCTGCAGGTG 509
QY 141 IysGlnIAspLeuAlaGlnGlyArgLeuThrCysAspAspThrSerAlaAlaLeuLeuIle 160
Db 510 AACAGAGACTGGCTCAAGGAGAGTGAAGTGAATGACACAGCCGAGCTCTTGATT 569
QY 161 SerHLeIleValGlnSerGluIleGlyAspPheAspGluAlaLeuAspArgGluHLeu 180
Db 570 TCACACTTGTGCATCTGAGATTGGGGAATTTGATGATACCTTGACAGAGAGCACTTA 629
QY 181 AlaLYsAsnLYsTYrIleProGlnGlnAspAlaLeuGlnAspLYsIleValGluPheHLe 200
Db 630 GCAAAAATATTAATACATACCTCAGCAAGACGCACTAGAGCAAAATCGTGAATTTTCA 689
QY 201 HLeAspHLeIleGlyGlnThrProAlaGlnSerAspPheGlnLeuLeuGluIleAlaArg 220
Db 690 CATACACACTTGACAAACACACAGAGATCAAGATTTCCAGCTCCAGAGATTGCCGT 749
QY 221 ArgLeuGlnMetTYrGlyIleArgLeuHLeProAlaLYsAspArgGluGlyThrLYsIle 240
Db 750 CGGCTAGAGATGATGTGAATCCGGTTCACCCGGCCAGAGACAGGAGGCAAGAGATC 809
QY 241 AsnLeuAlaValAlaAsnThrGlyIleLeuValPheGlnGlyPheThrLYsIleAsnAla 260
Db 810 AATCTGGCCCTTGCCAAACAGGGAATCTAGTGTTCAGGGGTTCATAAGATCAATGCC 869
QY 261 PheAsnThrAlaLYsValArgLYsLeuSerPheLYsArgLYsArgPheLeuIleLYsLeu 280
Db 870 TTCACTGGGCGCAAGGCGGAAGCTGAGCTCAAGAGGAAGCGCTTTTCATCAAGCTC 929
QY 281 ArgProAspAlaAsnSerAlaTYrGlnAspThrLeuGlnPheLeuMetAlaSerArgAsp 300
Db 930 CGGCGAGATGCCAATAGTGGCTACAGAGATCCTTGGAATTCCTGATGGCCATCCGGAT 989
QY 301 PheCysLYsSerPheTYrLYsIleCysValGlnHLeHLeAspPheArgLeuPheGlu 320
Db 990 TTCTGAAAGCTCTTGAAATCTGTGTGAACATCATGCTCTTTAGACTTTTGGAA 1049
QY 321 GluProLYsProLYsProLYsProValLeuPheSerArgLYsSerPheArgPheSer 340
Db 1050 GAGCCCAACCAAAAGCCCAAGCCGCTCTTTAAGCCGGGGGTCAATCATTTCCGTTCACT 1109
QY 341 GlyArgThrGlnLYsGlnValLeuAspTYrValLYsGlnGlyLYsIleLYsLYsValGln 360
Db 1110 GGTGCGACTCAGAGAGCGAGTCTCGACTATGTTAAAGAGAGAGACTAAAGAGGTGCGAG 1169
QY 361 PheGlnArgLYsHLeSerLYsIleHLeSerIleArgSerLeuAlaSerGlnProThrGlu 380
Db 1170 TTTGAAGAGAGACAGCAAGATTCAATTCGAGCCCTTGCTTCAACGCTTACAGAA 1229
QY 381 LeuAsnSerGlnValLeuGlnGlnSerGlnGlnSerThrSerLeuThrPheGlyGly 400
Db 1230 CTGAATTCGAAATGCTGAGAGCTCGACAGACAGCACAGCTTACATTTGAGAAAGGT 1289
QY 401 AlaGlnSerProGlyGlnSerCysArgArgLYsLYsGlnProLYsValSerAlaGly 420
Db 1290 GCGGAATCTCCAGGGGGCCAGAGCTGCGGCGAGAGAAAGAACGGAAGTTTCCGCGGG 1349
QY 421 GluProGlySerHLeProSerProAlaProArgArgSerProAlaGlyAsnLYsGlnAla 440
Db 1350 GAGCCCGGGGTGCGACCGAGCTCTGCGCGCAGAGAGAAAGCCCGCGGTAAACAAGCGGCG 1409
QY 441 AspGlyAlaAlaSerAlaProThrGlnGlnGlyGlnGlnValValLYsAspArgThrGln 460
Db 1410 GACGAGAGCGCTCGGGGCCACGAGAGAAAGAGAGGTGCTTAAGATTAGAACCCAG 1469

QY 461 GlnSerLYsProGlnProProGlnProSerThrGlySerLeuThrGlySerProHLeu 480
Db 1470 CAGAGTAACTCTAGGCCCCCGAGGCCAAGACAGAGTCTCTGACTGGAGTCTTCACTT 1529
QY 481 SerGlnLeuSerValAsnSerGlnGlyValAlaProAlaAsnValThrLeuSerPro 500
Db 1530 TCCGAGCTGTCTGTAACTCGCAGGGGGAGGTGGCCCTGCGCAACGTGACTGTCTCCC 1589
QY 501 AsnLeuSerProAspAspThrLYsGlnAlaSerProLeuIleSerProLeuAsnAspGln 520
Db 1590 AACCTAGCCCCGACACCAAGAGGCTCTCTCTATACAGCCGCTGTAATGACCAAG 1649
QY 521 AlaCysProArgThrAspArgGluAspGlnGlyArgArgLYsAspPheProThrAspLYs 540
Db 1650 GCCTGCCCGCGAGCGAGCATGAGTAGAGGGCCGAGAGAAAGATTCCCAACTGATAAA 1709
QY 541 AlaTYrPheIleAlaLYsGlnValSerThrThrGlnArgThrTYrLeuLYsAspLeuGlu 560
Db 1710 GCGTACTTCAATAGCTAAGAGAGTGTCTACACCGAGCGCAATATCTGAAGATCTCGAA 1769
QY 561 ValIleThrSerThrPheGlnSerThrValSerLYsGluAspAlaMetProGluAlaLeu 580
Db 1770 GTTATCACTTCGTGTTTCAAGACAGATGAGCAAAAGAGAGACCATGCGGAAAGACTG 1829
QY 581 LYsSerLeuIlePheProAspPheGlnProLeuHLeLYsPheHLeThrAspPheLeuLYs 600
Db 1830 AAAAGTCTATATTCGGAATTTTGAACCTTTGCAAAATTTCAATATTTTCCAAAG 1889
QY 601 GluIleGlnGlnArgLeuAlaLeuTYrGlnGlyArgSerAsnAlaGlnIleArgAspTYr 620
Db 1890 GAATTTGACCAACGACTTCCCTGTGGGAAGCGCTCAAAATGCCCAATACAGATTTC 1949
QY 621 GlnArgIleGlyAspValMetLeuLYsAsnIleGlnGlyMetLYsHLeuAlaAlaHLe 640
Db 1950 CAAGATTCGGGATTCATGCTGAAAGACATTCAGGAGATGAAGACCTGGCGGCTCAC 2009
QY 641 LeuTYrLYsHLeSerGlnAlaLeuGlnAlaLeuGlnAsnGlyIleLYsSerSerArgArg 660
Db 2010 CTGTGAAACACAGCAGAGCCTTGAGGCCCTGGAAGATGGAATCAAGAGCTCCGGCGG 2069
QY 661 LeuGlnAspPheCysArgAspPheGlnLeuGlnLYsValCysTYrLeuProLeuAsnThr 680
Db 2070 CTGGAAATCTTGCAGAACTTTGAGCTGCAAGAGTGTGTTACCTTACCGCTCAACACC 2129
QY 681 PheLeuLeuArgProLeuHLeArgLeuMetHLeTYrLYsGlnValLeuGlnArgLeuCys 700
Db 2130 TTCTCTCTCGGCGCACTGACCGGCTCATGCACTACACAGAGTCTTGAGCGGCTGTGC 2189
QY 701 LYsHLeHLeProProSerHLeAlaAspPheArgAspCysArgAlaAlaLeuAlaGluIle 720
Db 2190 AAACACACCCCGCGAGCCAGCCGACTTCAGGGACTGCCAGCCGCTTTGGCAGAGATC 2249
QY 721 ThrGlnMetValAlaGlnLeuHLeGlyThrMetIleLYsMetGlnAsnPheGlnLYsLeu 740
Db 2250 ACGGAATGATGGCAAGCTCCAGCGTACGATCAAGATGAAGAAATTTCCAGAAAGCTG 2309
QY 741 HLeGlnLeuLYsAspLeuIleGlyIleAspAsnLeuValAlaProGlyArgGluPhe 760
Db 2310 CACGAATCAAGAAAGATTGATGGCATTTGACATCTGTGTCTCCGGAAGGAGGCTTC 2369
QY 761 IleArgLeuGlySerLeuSerLYsLeuSerGlyLYsGlyLeuGlnGlnArgMetPhePhe 780
Db 2370 ATCCGTCTGGGACGCTTCAAGAGCTCGGGGAAGGGGCTCCAGACGCGCATGTTCTTC 2429
QY 781 LeuPheAsnAspValLeuLeuTYrThrSerArgLYsLeuThrAlaSerExtGlnPheLYs 800
Db 2430 CTGTTCAAGAGAGTCTGTTATACAGAGCCGGGGCTACAGCGCTTCAATCAAGTTTAAA 2489
QY 801 ValHLeGlyGlnLeuProLeuTYrGlyMetThrIleGlnGlySerGlnAspGluTYrGly 820
Db 2490 GTCCACGGGAGACTCCCGCTCATGTGCAATGAGATTCAGAGAGCGAAGACAGTGGGG 2549
QY 821 ValProHLeCysLeuThrLeuArgGlyGlnArgGlnSerIleIleValAlaAspSer 840

```

Db      2550 GTGCCCCACCTCCGACCTCCGGGGCCAGCGCATCATCTGCGCCGCACTTCT 2609
Qy      841 ArgSerGluMetGluValTyrValGluAapIleGlnMetAlaIleAapLeuAlaGluLys 860
Db      2610 CGGTCCCAAGATGAGAAAGTGCGTTGAGAGACATCCAGATGCGCATTTGCTGGCGAGAAAG 2669
Qy      861 SerSerSerProAlaProGluPheLeuAlaSerSerProProAlaAapLysSerProAap 880
Db      2670 AGCAGCAGCCCCGCGCTGAGATCTCTGGCCAGACGCCCCCTGACAAACATCTCCCTGAT 2729
Qy      881 GluAlaThrAlaAlaAapGlnGlnSerGluAapAapLeuSerAlaSerArgThrSerLeu 900
Db      2730 GAAGCCACCGCGCTGACAGAGATCAAGAGATCACTGACCGCTCGCGCACTCGCTG 2789
Qy      901 GluArgGlnAlaProAlaArgGlyAapThrMetValHisValCysThrPheAapAapThr 920
Db      2790 GAGCGCCAGCGCCCGCCAGCCGCGCAACAATGGTGCGCTGTCTGGCACCGCAACACC 2849
Qy      921 SerAlaSerMetValAapPheSerIleAlaValGluAapGlnLeuSerGlyAapLeuLeu 940
Db      2850 AGCGTCTCCATGTGTGAGACTTACGATCGCAGTGAAGATCATGTTGTCTGAAACCTGTG 2909
Qy      941 ArgLysPheLysAapAapSerAapGlyTyrGlnLysLeuTyrValValPheThrAapPheCys 960
Db      2910 AGGAAATTCACAAAACAGCAACGGGTGCGAGAACTGTGGTGTCTTACAAACTTCTGC 2969
Qy      961 LeuPhePheTyrLysSerHisGlnAapAapHisAapLeuAlaSerLeuProLeuLeuGly 980
Db      2970 CTGTCTTCTTCAATATCACACAGACATATCATCCCTTGGCCAGCTGCTGCTGCGGC 3029
Qy      981 TyrSerLeuThrIleProSerGluSerGluAapIleGlnLysAapTyrValPheLysLeu 1000
Db      3030 TACTCGCTCACACATCCCTCTGAGTCCGAGAACATCCAGAAAGACTACGTTCAACCTG 3089
Qy      1001 HisPheLysSerHisValTyrTyrPheArgAlaGluSerGlyTyrThrPheGluArgTyr 1020
Db      3090 CACTTCAAGTCCACGCTTACTTACTTCAAGGCGGAAAGCGAGTACACGTTGAAAGGTGG 3149
Qy      1021 MetGluValIleArgSerAlaThrSerSerAlaSerArgProHisValLeuSerHisLys 1040
Db      3150 ATGGAATGTAGTCCGACAGTGCACACAGCTCTGCTGCGACCCACGTTGATGATCACAAA 3209
Qy      1041 GluSerLeuValTyr 1045
Db      3210 GAGTCTTGTGTAT 3224

```

```

RESULT 6
BC071592      4932 bp      mRNA      linear      PRI 28-JUN-2005
LOCUS      BC071592
DEFINITION      Homo sapiens FERM, RhocGF (ARHGFR) and plectrokin domain protein 1
                (chondrocyte-derived), mRNA (cDNA clone MGC:87400 IMAGE:30342239),
                complete cds.
ACCESSION      BC071592
VERSION      BC071592.1 GI:47940451
SOURCE      MGC.
ORGANISM      Homo sapiens (human)
                Homo sapiens
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
                Homiidae; Homo.
REFERENCE      1 (bases 1 to 4932)
AUTHORS      Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
                Klauener,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
                Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
                Hopkins,B.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Heideh,F.,
                Diatchenko,L., Marusik,K., Farmer,A.A., Rubin,G.M., Hong,L.,
                Spatleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
                Scheetz,T.E., Brownstein,M.J., Uedin,T.B., Toshiyuki,S.,
                Abramson,J.P., Franke,C., Kane,S.S., Loquellano,N.A., Peters,G.J.,
                Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
                McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
                Morley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,

```

FEATURES

source

gene

CDS

Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Kettelman,M., Madan,A., Rodriguez,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,D., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalins,D.E., Scherch,A., Schein,J.E., Jones,S.J. and Marz,M.A.

Mammalian Gene Collection Program Team

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

2 (bases 1 to 4932)

NIH MGC Project

Direct Submission

Submitted (01-JUN-2004) National Institutes of Health, Mammalian Gene Collection (MGC), Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.ncl.nih.gov>

Contact: MGC help desk

Email: cgabs-remail.nih.gov

Tissue Procurement: Dr. Stefan Hansson

CDNA Library Preparation: Michael Brownstein / Ted Usdin

Laboratory

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305

Web site: <http://www.shgc.stanford.edu>

Contact: (Dickson, Mark) mcd@paxil.stanford.edu

Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>

Series: IRAX Plate: 168 Row: 1 Column: 8

This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 48928036.

Location/Qualifiers

1..4932

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="MGC:87400 IMAGE:30342239"

/tissue_type="Placenta, normal"

/clone_lib="NIH_MGC_147"

/lab_host="DH10B"

/note="Vector: pBluescriptR"

1..4932

/gene="FARP1"

/note="synonyms: CDEP, MGC87400, PLEKHC2"

/db_xref="GeneID:10160"

/db_xref="MIM:602654"

240..3470

/gene="FARP1"

/codon_start=1

/product="FARP1 protein"

/protein_id="AAH1592.1"

/db_xref="GI:47940452"

/db_xref="GeneID:10160"

/translation="MGEBORPTGSRNLAGPENSGLTLEGGOKPPPTSGKLVSITQMDIQDQAEVQPARQKTLDAVCNHNLYGQDYGLEPDPDKITWDLIDLPVQKIRPRHVVAVKVFPPDPDLOLBELTRYFLAIOVKDOLAQGRITCDTSAALLISHVQSEIGDPAFLDREHLAKKTYIPQDALLDKLVEFPHNHTGTPASDPOLLELRLKLEWGRGALHAKDREGRKIMLAANGCILVFGPTKIMAFNAKVRGLSPKRRFLIKLRPDANSAVYDTEFLASDRDPCKSEFKICVEHHAFLRLEEPKPKPVLFSGSSRFSGRTKOVLVDYKCGKVKVQFERKSHISIRLSAQPTLENSVLESOOS TSLRFGKASSPCGSCGCKEPEKVAAGEGSHSPAPRSPAGNCOADSAAPPEEVEVDKTRQSKRPQRPQPESTLTSPPHSELSSVNSQGVAAANTLSPNSPDKOASPLISPLNDQACPTDDDEGRKRKPTFDKAYFAKEVSTTERTYLQLEVTYSFQSVSKEDAPAEALKSLIFENPEFLKPHFNFKETLQRLALWEGSSNAQIRUYRYS DVMKLIQCKHKLAAHIMKXSEALALENLIKSRLENCRDFELQKVCYLPLKPTLRLPLRLHMYKQYLERLCKHPSHADFRCCRAALAEITEMAQLHGTMIKMENFQL

HELDKDLIGIDNLVPEGRPGSFSLMRTPLGLQARRIPCAERRRLLVKFEIRLSLS
KLSGKLGQRMFPLFNDVLLYTSRGLTASQFVHGOLPLYGCTTESSEBEMVPCU
TLRGOSITVVASRSEMEKXVEDIOMADLAEKSSPSSPNDKSPDPAAT
ADQESDDISASRTSLERODPIRGNTMHWCHRRRTSYSMVDPSTAVENQSLNLR
KPFNSNGWQCLWVFTNFCLEFYKSHODNHPNLASPLGSLTPESENIQDYVFK
LHKSHVYFRABSEYTFERMEVIRSATSSASRPVLSHKESTLVY"

ORIGIN

Alignment Scores:

Pred. No.:	0	Length:	4932
Score:	5431.50	Matches:	1044
Percent Similarity:	97.1%	Conservative:	0
Beet Local Similarity:	97.0%	Mismatches:	0
Query Match:	99.4%	Indels:	31
DB:	5	Gaps:	1

US-09-555-342b-2 (1-1045) x BC071592 (1-4932)

Qy 1 MetGlyGluIleGluGlnArgProThrProGlySerArgLeuGlyAlaProGluAenSer 20
Db 240 ATGGGAGAAATAGACAGAGGCGGACCCAGATCAGACTGGGGGCGCCGAGAAATTCG 239
Qy 21 GlyIleSerThrLeuGluArgGlyGlnLysProProThrProSerGlyLysLeuVal 40
Db 300 GGGATCAGTACTTGGAACTGGACAGAAAGCCGCCCAACCTTCAGGAAAACTCGTG 359
Qy 41 SerIleLysIleGlnMetLeuAspAspThrGlnGluAlaPheGlnValProGlnArgAla 60
Db 360 TCCATCAAAATCCAGAGTGTGAGATGACACAGAGAGCATTTGAAAGTTCCAAAGAGCT 419
Qy 61 ProGlyLysValLeuLeuAspAlaValCyAsenHisLeuAenLeuValGluGlyAspTyr 80
Db 420 CTGGGAGAGTGTGCTGGATGAGATGTCACCACTCAACCTCGTGGAAAGTGACTAT 479
Qy 81 PheGlyLeuGluIlePheProAspHisLysLysIleThrValIlePheLeuLeuLysPro 100
Db 480 TTGGGCTCAGATTTCTGTATCAAAAGATCAGCGTGTGGCTGATCTCTAAAGCC 539
Qy 101 IleValIleGlnIleArgArgProLysHisValValLysPheValValLysPhePhe 120
Db 540 ATTGTGAACAGATTAGAGGCCAAGACGTTGTGTTAAAGTTGTGGTGAATTTCTTT 599
Qy 121 ProProAspHisThrGlnLeuGlnGlnGluLeuThrArgTyrLeuPheAlaLeuGlnVal 140
Db 600 CGCGCTGACCAACAACACTCCAAGAAAGACTCAAGAGTACCTGTGCGGTGCAAGGTG 659
Qy 141 LysGlnAspLeuAlaGlnGlyArgLeuThrCyAsenAspHisSerAlaAlaLeuLysIle 160
Db 660 AAGCAGGACTTGGCTCAAGGCAAGTTGACGTGTAAATGACACAGCGAGCTCTTGATT 719
Qy 161 SerHisIleValGlnSerGluIleGlyAspPheAspGluAlaLeuAspArgGluHisLeu 180
Db 720 TCCACACTTTGCAATTTGAGATTTGGGGAATTTGATGAACCTTTGACAGAGAGCACTTA 779
Qy 181 AlaLysAsnLysTyrIleProGlnGlnAspAlaLeuGluAspLysIleValGluPheHis 200
Db 780 GCAAAAATAATTAATACATCACTCAGCAAGACGCACTAGAGCAAAATCGTGAATTTCCAC 839
Qy 201 HisAsenHisIleGlyGlnThrProAlaGluSerAspPheGlnLeuLeuGluIleAlaArg 220
Db 840 CAATACCACTTGGACAAACACAGCAGATCAGATTTCCAGCTCTCAGAGATTCGCCGT 899
Qy 221 ArgLeuGluMetTyrGlyIleArgLeuHisProAlaLysAspArgGluGlyThrLysIle 240
Db 900 CGGCTGAGATGATGAGATCCGGTTGACCCGGCCAAAGACAGGAGAGGCAAGAGATC 959
Qy 241 AsnLeuAlaValAlaAsenThrGlyIleLeuValPheGlnGlyPheThrLysIleAsnAla 260
Db 960 AATCTGGCCGTGGCAACAGGGAATCTAGTGTTCAGGGTTTCACATAAGATCAATGCC 1019
Qy 261 PheAsnThrAlaLysValArgLysLeuSerPheLysArgLysArgPheLeuIleLysLeu 280
Db 1020 TTCAACTGGGCAAGGTGGAGGTGAGCTTCAAGAGAAAGCGCTTTCATCAAGCTC 1079

Qy 281 ArgProAspAlaAsnSerAlaTyrGlnAspThrLeuGluPheLeuMetAlaSerArgAsp 300
Db 1080 CGGCGAGATGCCAATAGTGTGACGACGATACCTTGAATTCCTGATGGCCACTCGGAT 1139
Qy 301 PheCyAlaSerPheThrLysIleCyValGluHisAlaPhePheArgLeuPheGlu 320
Db 1140 TTCTGCAAGTCTTCTGAAAAATCTGTGTGAACATCATGCTTCTTAGACTTTTGGAA 1199
Qy 321 GluProLysProLysProLysProValLeuPheSerArgGlySerSerPheArgPheSer 340
Db 1200 GAAGCCAAACCAAGGCCCAAGCCGCTCTTGAAGCCGGGGGTCAATCATTTCCGTTCACT 1259
Qy 341 GlyArgThrGlnLysGlnValLeuAspTyrValLysGluGlyGlyHisLysLysValGln 360
Db 1260 GGTGGACCTCAGAGAGAGGTTCTCGACTATGTTAAAGAGAGAGACATAAGAGGTGCAG 1319
Qy 361 PheGluArgLysHisSerLysIleHisSerIleArgSerLeuAlaSerGlnProThrGlu 380
Db 1320 TTGAAAGAGAGACAGCAAGATTCATTCCTGAGACCTTGCTTACAGCCTACAGAA 1379
Qy 381 LeuAsnSerGluValLeuGluGlnSerGlnGlnSerThrSerLeuThrPheGlyGluGly 400
Db 1380 CTGAATTCGAAAGTGTGAGACAGTCTCAGACAGACACAGCTTACATTTGAGAAAGT 1439
Qy 401 AlaGluSerProGlyGlyGlnSerCyAsArgGlyLysGluProLysValSerAlaGly 420
Db 1440 GCGGATCTCCAGGGGGCCAGAGCTGCCGCGAGAGAAAGAACCAAGGTTTCCGCCGG 1499
Qy 421 GluProGlySerHisProSerProAlaProArgArgSerProAlaGlyAsnLysGlnAla 440
Db 1500 GAGCCGGGGTGGCAGCCGAGCTGTGCGAGAGAGAGAGCCCGCGGTAAACAGAGCGCG 1559
Qy 441 AspGlyAlaAlaSerAlaProThrGlnGluGluGluValValLysAspArgThrGln 460
Db 1560 GACGGAGCCGCTCGGCGCCACGAGAGAGAGAGAGAGTCTTAAAGATAGAGACCCAG 1619
Qy 461 GlnSerLysProGlnProProGlnProSerThrGlySerLeuThrGlySerProHisLeu 480
Db 1620 CAGAGTAAACCTAGCCCGCCGAGCCCAACAGCAGCTCCCTGACTGCGAGTCTCACTT 1679
Qy 481 SerGlnLeuSerValAsnSerGlnGlyValAlaProAlaAsnValThrLeuSerPro 500
Db 1680 TCCGAGCTGTGTGAATCGCAGGGGGAGAGTGGCCCTGCGCAACCTGACTGTCTCCC 1739
Qy 501 AsnLeuSerProAspThrLysGlnAlaSerProLeuIleSerProLeuAsnAspGln 520
Db 1740 AACCTAGCCCGACACCAAGAGGCTCTCCCTTATCAGCCCGCTGTAATGACAG 1799
Qy 521 AlaCyProArgThrAspAspGluAspGluGlyArgArgLysArgPheProThrAspLys 540
Db 1800 GCCTGCCCGGAGCGAGCATGAGTGAAGGGCCGAGAGAAAGATTTCCCACTGATTA 1859
Qy 541 AlaTyrPheIleAlaLysGluValSerThrThrGluArgThrTyrLeuLysAspLeuGlu 560
Db 1860 GCCTACTCATAGCTAAGAGAGTGTCTACACCGAGCAACATATCTGAAGATCTCGAA 1919
Qy 561 ValIleThrSerThrPheGlnSerThrValSerLysGluAspAlaMetProGluAlaLeu 580
Db 1920 GTTATCACTTCGGGTTTCAAGACAGATGAGCAAGAGAGAGAGCCCATGCGGAGACCTG 1979
Qy 581 LysSerLeuIlePheProAspPheGluProLeuHisLysPheHisThrAsnPheLeuLys 600
Db 1980 AAAAGTCAATATTCGGAATTTTGAACCTTTGCAAAATTTCAATCAATTTTCCAAG 2039
Qy 601 GluIleGluGlnArgLeuAlaLeuThrGluGlyArgSerAsnAlaGlnIleArgAspTyr 620
Db 2040 GAAATTAGACCAACGATTCCTCGTGGAGAGGCGCTCAAAATGCCAAATACAGAGTTAC 2099
Qy 621 GlnArgIleGlyAspValMetLeuLysAsnIleGlnGlyMetLysHisLeuAlaHis 640
Db 2100 CAAGATTCGGGAGTATCTAGCTGAAGAACATTCAGGGCATGAAGCACTGGGCGCTCAC 2159

```

Qy 641 LeuTrpHisSerGluAlaLeuGluValLeuGlnIleuSerSerArgArg 660
Db 2160 CTGGAGAGTACACGAGGCTTGGAGGCTTGGAGGATGAAATGAAAGAGTCCCGGCGG 2219
Qy 661 LeuGluAsnProPheCysArgAspPheGluLeuGlnIleuValCysTyrLeuProLeuAsnThr 680
Db 2220 CTGGAGAACTTCTCAGAGACTTTCAGCTGCGAAGAGGTGTGTCTTCAACGCTCAACGCC 2279
Qy 681 PheLeuLeuArgProLeuHisArgLeuMetHisTyrIleuValIleuGluArgLeuCys 700
Db 2280 TTCTCTCTGGGGCCACCTGCACCGGCTCATGCTACAGCAGCGCTCTGGAGCGGCTGTGC 2339
Qy 701 LysHisAspProSerHisAlaAspPheArgAspCysArgAlaAlaLeuAlaGluIle 720
Db 2340 AATACACACCGCGGAGCCGACGCGGACTTCAGGAGCTGCCAGCGCTTGGCAGAGATC 2399
Qy 721 ThrGluMetValAlaGlnIleuHisGlyThrMetIleuSerGluAsnProGlnIleu 740
Db 2400 ACGGAGATGGTGGACAGCTCCACGATACGATGATCAAGATGAGAAATTTCCAGAGCTG 2459
Qy 741 HisGluLeuIleuValAspLeuIleGlyIleAspAsnLeuValIleProGlyArg----- 758
Db 2460 CACGAACTCAAGAAAGATTGATTGCGCATTCCTTGCTTCCGGGAGGCGCTGGC 2519
Qy 758 ----- 758
Db 2520 TCCCTTTCCCTTATAGAGACTCCACACTTATGAGACAGCGGCGCATCCCATGTGCTCCA 2579
Qy 759 -----GluPheIleArgLeuGlySerLeuSerIleu 769
Db 2580 GAGAGAGGCGCATGTTATTAGTAAGAGATTTCATCCGTGGGACAGCTCAGACAGCTC 2639
Qy 770 SerGlyIleuGlnIleuGlnIleuArgMetPhePheLeuPheAsnAspValLeuLeuTyrThr 789
Db 2640 TCGGGGAGGGGCTCCAGACGCGCATGTCTTCTGTTCACAGCGCTCTGTATTAACG 2699
Qy 790 SerArgGlyLeuThrAlaSerAsnGlnPheValHisGlyIleuLeuProLeuTyrGly 809
Db 2700 AGCCGGGGGCTGACGCGCTCCATCATGTTTAAATCCAGGGGCGCTCCGCTTATGGC 2759
Qy 810 MetThrIleuGlnIleuSerGluAspGluTyrGlyValProHisCysValLeuThrLeuArgGly 829
Db 2760 ATGACGATTAGAGAGAGCAGAGAGAGTGGGGGTGCCCATCTCGACCTCGGGGC 2819
Qy 830 GlnArgGlnSerIleIleValAlaIleSerSerArgSerGluMetGluIleTyrValGlu 849
Db 2820 CAGCGGCGATCCATCATCTGTGGCCGCGCATGTTCTCGTCCGAGATGAGAGTGGGTGAG 2879
Qy 850 AspIleGlnMetAlaIleAspLeuAlaGluIleSerSerSerProAlaProGluPheLeu 869
Db 2880 GACATCCAGATGGCATTTGACTGCGAGAGAGAGCAGCGGCGGCTTGAATCTCTG 2939
Qy 870 AlaSerSerProAspAsnIleSerSerProAspGluAlaThrAlaAlaAspGlnIleuSer 889
Db 2940 GCCAGCAGCCGCCCTGACAAACATCCCTGATGAAGCAGCCGCGGTGACAGAGATCA 2999
Qy 890 GluAspAspLeuSerAlaSerArgThrSerLeuGluArgGlnAlaProHisArgGlyAsn 909
Db 3000 GAGGATACCTGAGCGGCTCGCGCACATCGCTGAGGCGCCGCGCGCCGACCGGCGAAC 3059
Qy 910 ThrMetValHisValCysTrpHisArgAsnThrSerValSerMetValAspPheSerIle 929
Db 3060 ACAATGGTGCACGTGTCTGGCAGCAGCAGCGCTTCATATGTGAGACTTCAGCATC 3119
Qy 930 AlaValGluAsnGlnIleuSerGlyAsnLeuLeuArgIleuPheIleAsnSerAsnGlyTyr 949
Db 3120 GCACTGAGATCAGTGTGTGGAACCTGTGAGGAATTCAAAACAGCAACGCGGTGG 3179
Qy 950 GlnIleuLeuTrpValIlePheThrAsnProPheCysLeuLeuPheTyrIleSerHisGlnAsp 969
Db 3180 CAGAAAGTGTGGGTGTTCACAAACTTGTGCTGTCTTCAAAATCACACAGAGAC 3239
Qy 970 AsnHisProLeuAlaSerLeuProLeuGlnIleTyrSerLeuThrIleProSerGluSer 989

```

```

Db 3240 AATATCCCTTGGCAGCTCGCTCTGCTGCTACCTGCTACCAATCCCTCGAATCC 3299
Qy 990 GluAsnIleGlnIleuAspTyrValPheIleuHisPheIleuSerHisValIleTyrPhe 1009
Db 3300 GAGAACATCCAGAAAGACTAGTGTTCAGCTGACACTTCAAGTCCACGCTCTACTTTC 3359
Qy 1010 ArgAlaGluSerGlnIleuTyrTrpPheGluArgTrpMetGluValIleAspSerAlaThrSer 1029
Db 3360 AGGCGGAGAAAGCAGATACAGTTCGAAAGGTGATGAGATGATCCGACGTGCCACAGC 3419
Qy 1030 SerAlaSerArgProHisValLeuSerHisIleGluSerLeuValIleTyr 1045
Db 3420 TCTGCTCGGAGACCCACGTTGATGATCAAAAGATCTTGTGTAT 3467

RESULT 7
LOCUS CO724059 2937 bp DNA linear PAT 03-FEB-2004
DEFINITION Sequence 9993 from Patent WO02068579.
ACCESSION CO724059
VERSION CO724059.1 GI:42284916
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Homo sapiens
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominae; Homo.
REFERENCE
1 Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
Kits, such as nucleic acid arrays, comprising a majority of
humanexons or transcripts, for detecting expression and other uses
thereof
Patent: WO 02068579-A 9993 06-SEP-2002.
JOURNAL
FEATURES
source 1..2937
location/Qualifiers
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ALIGNMENT Scores:
Pred. No.: 0 Length: 2937
Score: 4095.50 Matches: 825
Percent Similarity: 79.3% Conservative: 4
Best Local Similarity: 78.9% Mismatches: 32
Query Match: 75.0% Indels: 185
DB: Gaps: 7

US-09-555-342B-2 (1-1045) x CO724059 (1-2937)
Qy 58 GlnArgAlaProGlyIleuValLeuLeuAspAlaValCysAsnHisIleuAsnLeuValGlu 77
Db 1 CAAGAGAGCTCTGGAGAGGTGCTGCTGATGATGATTTGCAACCACTCACTGCTGAA 60
Qy 78 GlyAspTyrPheGlyIleuGluPheProAspHisIleuValIleThrValIlePheAspLeu 97
Db 61 GGTGACTATTGTGCTCGAGTTCCTGATCAAAAAGATCAG----- 105
Qy 98 LeuIleProIleValIleGlnIleArgArgProIleHisIleValIleIlePheValIle 117
Db 106 -----GCCAGGTTCCCGGTG 120
Qy 118 LysPhePheProProAspHisThrGlnIleuGlnGlnIleuThrArgTyrLeuPheAla 137
Db 121 ACCTCTGCGCTCTTGGCTGAGCGCCAGCTCTGTGATCTTCGACGTACCTGTGCGG 180
Qy 138 LeuGlnValIleGlnAspLeuAlaGlnIleuValArgLeuThrCysAsnAspThrSerAlaIle 157
Db 181 CTGAGGTGAAGAGCACTTGGCTCAAGGCGAGGTGACGTATATGACACAGCGAGCT 240
Qy 158 LeuLeuIleSerHisIleValIleGlnSerGluIleGlyAspPheAspGluAlaLeuAspArg 177

```

Db 241 CTCCTGATTTACACACTTGTCGAATCTGAGATTGGGGATTTTGATGAGCCCTTGCACAGA 300
QY 178 GlnH1eLeuAlaIysAsnIysTyrIleProGlnGlnAspAlaLeuGluAspIleVal 197
Db 301 GACACCTTACCAAAATAATATACATCTCTCAGCAGACCCACTTACGACAAATCTG 360
QY 198 GluPheH1eH1AspH1sIleGlyGlnThrProAlaGluSerAspPheGlnLeuGlu 217
Db 361 GAATTCACCACTAACCACTTGACAAACACGACAGAAATCAATTCAGCTCTCAGAG 420
QY 218 IleAlaArgArgLeuGluMetTyrGlyIleArgLeuH1sProAlaIysAspArgGluGly 237
Db 421 ATGGCCGTCGCTAGAGATGATGAGATCCGGTTCAGCCGCGCAAGACAGAGAGGC 480
QY 238 ThrIleI1eAsnLeuAlaValAlaAsnThrGlyIleLeuValPheGlnGlyPheThrIys 257
Db 481 ACBAATATCAATCTGGCCGCTTGCACACCGGAATTCAGTTCAGGGTTTCATACAG 540
QY 258 IleAsnAlaPheAsnThrAlaIysValArgIysLeuSerPheIysValArgPheLeu 277
Db 541 ATCAATGCTTCAACTGGCGCCAGAGTCCGAGAGCTGACTTCAAGAGAGCGCTTCTC 600
QY 278 IleIysLeuArgProAspAlaAsnSerAlaTyrGlnAspThrLeuGluPheLeuMetAla 297
Db 601 ATCAAGCTCGGCGCAGATGCAATAGTCAGTACAGGATTCCTGGAATTCAGATGGCC 660
QY 298 SerArgAspPheCysIysSerPheThrIleI1eCysValGluH1eH1sAlaPhePheArg 317
Db 661 AGTCGGGATTTCTGCAGATCTCTTGAAATCTGTGTGAACATCAATGCTCTTTAGA 720
QY 318 LeuPheGluGluProIysProIysProIysProValLeuPheSerArgIysSerPhe 337
Db 721 CTTTGGAGAGCCCAACCAAGCCCAAGCCGCTCTTTCAGCGGGGGTCAATATT 780
QY 338 ArgPheSerGlyArgThrGlnIysGlnValLeuAspTyrAllyIysGluGlyI1sIys 357
Db 781 CGGTTCAGGTGTGGACTCAGAAAGCAGGTTCTGCATATGTAAAGAGAGACAAAG 840
QY 358 LysValGlnPheGlnIuArgIysH1sSerIysI1eH1sSerI1eArgSerLeuAlaSerGln 377
Db 841 MAGGTGAGTTTAAAGAGACACAGCAATTCATTCTTCGAGACCTTGTCTTACAG 900
QY 378 ProThrGluLeuAsnSerGluValLeuGluGlnSerGlnI1sSerThrSerLeuThrPhe 397
Db 901 CCAACAGAACTGAATTCGAAAGTCTGAGACATCTCAGCAGACACACAGCCCTTAATTT 960
QY 398 GlyGluGlyAlaIuSerProGlyIysGlnSerCysArgArgGlyIysGluProIysVal 417
Db 961 GAGAAAGGTGCCAATCTCCAGGAGGCGCAGAGCTGCGCGCAGAGAAAGAAAGGTT 1020
QY 418 SerAlaGlyIuProGlySerH1sProSerProAlaProArgArgSerProAlaGlyAsn 437
Db 1021 TCGCGCGGGAGCGGGGTGCAACCGAGCTGCGCGCGAGAGAAAGCCCGGGGTAAAC 1080
QY 438 LysGlnAlaAspGlyAlaAlaSerAlaProThrGluGluGluGluValI1sIysAsp 457
Db 1081 AACCAAGCGGACGAGCGCTGCGCGCCCAACGAGAGAAAGAGAGGAGTCTTAAGGAT 1140
QY 458 ArgThrGlnGlnSerIysProGlnProProGlnProSerThr----- 471
Db 1141 AGGACCCAGAGAGTAACTCAGCCCCGACGCAAGACAGGTTCAGCATCCCGGGCT 1200
QY 471 ----- 471
Db 1201 GCCAGAGGACGACGATCTTCATTCCTTTCATTCGACGAGTACGTAAGTACGAATAC 1260
QY 471 ----- 471
Db 1261 GACCTTCTTAGAGAACAGCGTCCCGATCCCGATCCCGATCCCGATCCCGATCCCAAACA 1320
QY 472 -----GlySerLeuThrGlySerProH1sLeuSerGluLeuSerValAsn 486
Db 1321 AATGACATTTATGAGGCTCCCTGACTGGCAGTCTTCACTTTCGAGCTGTCTGTGAAC 1380

QY 487 SerGlnGlyIysValAlaProAlaAsnValThrLeuSerProAsnLeuSerProAspThr 506
Db 1381 TCGCAGGGGGAGGTGGCCCCCTGCAACGTGACTTGTCTCCAACTGAGAGCCGACACC 1440
QY 507 LysGlnAlaSerProLeuI1sSerProLeuLeuAsnAspGlnAlaCysProArgThrAsp 526
Db 1441 AAGCAGGCTCTCCCTTGTATCAGCCCGCTGTAAGTACAGAGGCTTGGCCCGAGAGGAC 1500
QY 527 AspGluAspGluGlyArgArgIysArgPheProThrAspIysValAlaThrPheI1eAlaIys 546
Db 1501 GATGAGATGAGGGCGCGAGAGAGATTCACCACTGATTAAGCGTATCTCATAGCTTAAG 1560
QY 547 GluValSerThrThrGluArgThrTyrLeuIysAspLeuGluValI1eThrSerThrPhe 566
Db 1561 GAAGTCTTACACCGAGGAAATATCTGAGAGATCTGAACTTCACTTGTGTGTTT 1620
QY 567 GlnSerThrValSerIysGluAspAlaMetProGluAlaLeuIysSerLeuI1ePhePro 586
Db 1621 CAGAGCAGATGACGAAAGAGACCGCCATGCCGAGACCTGMAAGTCTCATATTCGCG 1680
QY 587 AsnPheGluProLeuH1sIysPheH1sThrAsnPheLeuIysGluI1eGluGlnArgLeu 606
Db 1681 AATTTGAACCTTTGACAAATTTCTACTTAATTTCTCAAGGAAATTAAGCAACACTT 1740
QY 607 AlaLeuTrpGluGlyArgSerAsnI1eGlnI1eArgAspTyrGlnArgI1eGlyAspVal 626
Db 1741 GCCCTG----- 1746
QY 627 MetLeuIysAsnI1eGlnI1eMetIysH1sLeuAlaH1sLeuTrpIysH1sSerGlu 646
Db 1746 ----- 1746
QY 647 AlaLeuGluAlaLeuGluI1eGlnI1eIysSerSerArgArgLeuGluI1eAsnPheCysArg 666
Db 1747 -----TTC--- 1749
QY 667 AspPheGluLeuGluI1eIysValCysTyrLeuProLeuAsnThrPheLeuLeuArgProLeu 686
Db 1750 ----GTTGAGCTGACAGCTCTAGGGGAGAGACCCCTGCTGCTT----- 1791
QY 687 H1sArgLeuMetH1sTyrIysGlnValLeuGluI1eGluCysIysH1sH1sProProSer 706
Db 1792 -----CCGCTGCA 1800
QY 707 H1sAlaAspPheArgAspCysArgAlaAlaLeuAlaGluI1eThrGluMetValAlaGln 726
Db 1801 GACTGTGTATGCTTCTTCCACAGCCGCTTGGCAGAGATCACGAGATGTGTGGCAGAG 1860
QY 727 LeuH1eGlyThrMetI1eIysMetGluAsnPheGlnIysLeuH1eGluLeuIysIysAsp 746
Db 1861 CTCACGATGATGATCAAGATGAGAAATTCAGAAACCTGACAGCAACCTCAAGAAAGAT 1920
QY 747 LeuI1eGlyI1eAspAsnLeuValProGlyArgGluPheI1eArgLeuGlySerLeu 766
Db 1921 TTGATTTGGATTAACATTTGTGTGTTCCGGAAGGAAATTCATCCGTCTGGCAGCCTC 1980
QY 767 SerIysLeuSerGlyIysGlyLeuGlnGlnArgMetPhePheLeu----- 781
Db 1981 AGCAGCTCTCGGGAGAGGGGCTCCAGCAGCCGATTTCTTCTGTGATGAGAGAGAGC 2040
QY 781 ----- 781
Db 2041 GGCTGTCTCAACAGATTTGTGTCACTGGGCAAGCAAGCTCCCAAGGCGGAGCTCG 2100
QY 782 -----PheAsnAspValLeuLeuTyrThrSerArgGlyLeuThrAlaSerAsnGln 798
Db 2101 GCCACCAATTAAACGACGCTGTGTATACAGAGCGGGGGCTGACGCGCTTCATATCAG 2160
QY 799 PheIysValH1eGlyI1eLeuProLeuTyrGlyMetThrI1eGluI1eIysSerGluAspGlu 818
Db 2161 TTTAAAGTCCACGGGAGCTCCCGCTCTATGTGATGACGATTAAGAGAGAGAGAGAGAG 2220

ORLALWEGNSNAHVGRDYGRIQIDVMLKNIQGMKHLAHLMKHSEALEAETSIKGSR
LEHFCRDLFOKVCYLPLNTFLRLPLRLMHWYGLVERLCKHPPHNAFDRDRAALA
EIRWVAADLHGTQWIMKMNFOKHELKDKLDGINLVTPGEPFRLSLSLSTGKGO
RMFLPNDULYTSRGLTASNORKGOLPLVMTIERSRECVGHCLTRBORSI
IYAASSRSEWEKEMEDIQMAIDLAESKNGTPPBLASSPBDKSPDEATRAADESD
LSASRYSLEKQAPHRGNTVWVCWHRSTSVWDFSIIVENQSLMLRFPKSNQW
KLMVVTNFECLFPYKSHODSHPLASLPLGLVSTLTPSESNIKHDYVFKLHFKSHYV
FRASEYTERFERNMEIVIRSATSSASRAHILSHKESHLY"
999. .1556

misc_feature

misc_feature

misc_feature

/gene="Farp1"
/note="RhoGEF; Region: Guanine nucleotide exchange factor
for Rho/Rac/Cdc42-like GTPases"
/db_xref="CDD:smat00325"
1695. .1946
/gene="Farp1"
/note="PH; Region: Pleckstrin homology domain. Domain
commonly found in eukaryotic signalling proteins. The
domain family possesses multiple functions including the
abilities to bind inositol phosphates, and various
proteins. PH domains have been found to possess inserted
domains (such as in PLC gamma, syntrrophins) and to be
inserted within other domains. Mutations in Brutons
tyrosine kinase (Btk) within its PH domain cause X-linked
agammaglobulinaemia (XLA) in patients. Point mutations
cluster into the positively charged end of the molecule
around the predicted binding site for phosphatidylinositol
lipids"
/db_xref="CDD:smat00233"
2175. .2465
/gene="Farp1"
/note="PH; Region: Pleckstrin homology domain. Domain
commonly found in eukaryotic signalling proteins. The
domain family possesses multiple functions including the
abilities to bind inositol phosphates, and various
proteins. PH domains have been found to possess inserted
domains (such as in PLC gamma, syntrrophins) and to be
inserted within other domains. Mutations in Brutons
tyrosine kinase (Btk) within its PH domain cause X-linked
agammaglobulinaemia (XLA) in patients. Point mutations
cluster into the positively charged end of the molecule
around the predicted binding site for phosphatidylinositol
lipids"
/db_xref="CDD:smat00233"

ORIGIN

Alignment Scores:

Pred. No.: 9,39e-316 Length: 3902
Score: 3963.50 Matches: 758
Percent Similarity: 93.8% Conservative: 22
Best Local Similarity: 91.1% Mismatches: 49
Query Match: 72.6% Indels: 3
DB: 6 Gaps: 2

US-09-555-342B-2 (1-1045) x BC030329 (1-3902)

QY 217 G|U|I|E|A|A|A|G|A|G|L|E|U|G|I|U|E|T|Y|G|I|L|E|A|G|L|E|U|H|I|S|P|R|O|A|L|A|Y|S|A|P|R|G|L|U|216
|:::|
DB 12 GAGGTGGCCCGTGGCTTGAAGATGATCGAATCCGTTTACACCTCGCAAGACGAGGA 71
QY 237 G|Y|T|H|L|Y|E|L|E|A|N|L|E|U|A|A|V|A|L|A|E|N|T|H|G|I|Y|L|E|U|V|A|L|P|H|E|G|I|N|G|I|P|H|E|T|H|R|256
|:::|
DB 72 GGCACCAAGATCAACCTGCTGCGCAACACATGGAATTCAGTGTTCAGGCGCTTCACG 131
QY 257 L|Y|E|I|L|A|S|H|A|A|P|H|E|A|N|T|P|R|A|L|Y|E|V|A|L|A|G|L|E|U|S|E|R|P|H|E|Y|S|A|G|L|Y|S|A|G|P|H|E|276
|:::|
DB 132 AATATCAATGCTTTTAACTGGCCCAAGTGGAGAGTGAAGCTTCAAGCCGAAAGCGCTTT 191
QY 277 L|E|U|I|E|L|E|U|A|G|R|P|R|O|A|P|R|A|A|E|N|S|E|R|A|L|Y|T|G|I|N|A|S|P|H|R|E|U|G|I|U|P|H|E|L|E|U|E|T|296
|:::|
DB 192 CTTATCAAGCTGACGACTGACGTGAATGCTCATACGAGACACCTTGAGAGTTCTCGATG 251
QY 297 A|L|S|E|R|A|G|A|P|H|E|C|Y|L|Y|S|E|R|P|H|E|T|P|L|Y|E|L|C|Y|A|V|A|G|L|U|H|I|S|A|L|A|P|H|E|P|H|E|316
|:::|

DB 252 GCCGGCCGAGACTTTCGAAGCTTTCTGAAAAATATGTGTGAGACCAATGCTTCTTC 311
QY ArgLeuPheGluGluProLysProLysProLysProValLeuPheSerArgGlySerSer 336
317
DB 312 AGACTTTTGAGAGACCTTAAGCCAAAGCCAAAGCTGTTCTTTCAGCCGAGGTTGTTCC 371
QY PheArgPheSerGlyArgThrGlnLysGlnValLeuAspTyrValLysGluGlyYHs 356
337
DB 372 TTCGGTTTATGTGTGCACTCGAAGCAAGTTCTTGATATATGTAAAGAGCGGGCCAC 431
QY
QY 357 LysLysValGlnPheGluArgLysHsSerLysHsHsSerLLeArgSerLeuAlaSer 376
DB 432 AAGAAAGTCAAGTTTAAAGAAAGACACCAAAATCCATCCACCGAGGCTGTGCTCA 491
QY GlnProThrGluLeuAsnSerGluValLeuGluGlnSerGlnLysThrSerLeuThr 396
377
DB 492 CAGCCCAAGACACCGAATTCAGAAAGTCCAAAACAGTCTCCACAGAGCGCAAGCTTACG 551
QY PheGlyGluGlyAlaGluSerProGlyGlyGlnSerCysArgArgGlyLysGluProLys 416
397
DB 552 TTTGAGAAAGGTACGAGTCTCGGGGTGGCAGAGCTGCAGCAAGCGAAAGAAACAAA 611
QY PheGlyAlaGlyGluProGlySerHsProSerProAlaProArgSerProAlaGly 436
417
DB 612 GCCTGTACCTTGGAAGTTGGGACCTCAACAAAGCCCGCTCCCTAAGAGCCCTCCAGGG 671
QY AsnLysGlnAlaAspGlyAlaAlaSerAlaProThrGluGluGluValVal--- 455
437
DB 672 AGCAAAAGCTGCAGATGAACCAACCGTGTGCGCCTGAGAGAAAGAGAAAGAGGT 731
QY 456 ---LysAspArgThrGlnLysSerLysProGlnProProGlnProSerThrGlySerLeu 474
732
DB 732 GGCAGAGAGGATCCGGCAAGTAACTCCAGGCCCCCGCAGCAAGCAGGCTCCCTG 791
QY ThrGlySerProHsLeuSerGluLeuSerValAsnSerGlnGlyValAlaProAla 434
475
DB 792 ACTGTAAGCCCTCACTTCAGAGCTGTCTCATCACTCAAGAGAGAGAGCGGCCAGCC 851
QY AsnValThrLeuSerProAsnLeuSerProAspThrLysGlnAlaSerProLeuLys 514
495
DB 852 AACGTAACTGTCTTCCAACTGAATCTTGACAAAGCAGGCTCTCCCTGATCAGC 911
QY 515 ProLeuLeuAsnAspGlnAlaCysProArgThrAspAspGluAspGluArgArgLys 534
912
DB 912 CGGCTCCTGAAGACCAAGCGTGCACCAAGACCGAGCAGAGAAAGAGGCGCGCAAG 971
QY 535 ArgPheProThrAspLysAlaTyrPheLeuAlaLysGluValSerThrThrGluArgThr 554
535
DB 972 AGATTTCCAACGACAAAGCGTACTCATTTGCTTAAGAAAGTCTCCACACAGAGCGCACG 1031
QY TyrLeuLysAspLeuGluValIleThrSerTyrPheGlnSerThrValSerLysGluAsp 574
555
DB 1032 TACCTTAAGACCTTGAAGTCAATGCTTGTGTTTCAAGACACATCGACAAAGAGAC 1091
QY 575 AlaMetProGluAlaLeuLysSerLeuIlePheProAsnPheGluProLeuHsLysPhe 594
1092
DB 1092 TCCATCCCGGAAGCCTTGAAGAGTCTCATATTCCCAATTTTGAACCTTTGACCAAGTTTC 1151
QY 595 HisThrAsnPheLeuLysGluLleGluGlnArgLeuAlaLeuTyrGluValArgSerAsn 614
595
DB 1152 CATACCAATTTTCTGAAGAAATTGAGCAACGACTTGCCCTGAGGAAGGCCCTCAAAAT 1211
QY AlaGlnIleArg---AspTyrGlnArgGlyGlyAspValMetLeuLysAsnIleGlnGly 633
615
DB 1212 GCCCAAGTCAAGAGATTAACAAAGATTCGGGAGCTCATGCTGAAAAATTCAGGGC 1271
QY 634 MetLysHsLeuAlaAlaHsLeuTyrLysHsSerGlnAlaLeuGluAlaLeuGluAsn 653
1272
DB 1272 ATGAAGACCTAGGCGGCCACCTGTGAGAGCAGAGTGAAGGCGCTTGAGGCGCTGAGACC 1331
QY 654 GlyLleLysSerSerArgArgLeuGluAsnPheCysArgAspPheGluLeuGlnLysVal 673
1332
DB 1332 TCCATCAAGGCTCGGGGCGGTGGAACATTTCTGCCGAGACTTCGAGCTGCAGAAAGGTG 1391

QY 674 CyetYrleuProleuAenThrPheleuLeuArgleuMerHisETyLys 693
 Db 1392 TGTATCTGGCGGCTCAACACTTCTCTGCGCCGCGGCTGACATGACATTAADAG 1451
 QY 694 GlnValLeuGluArgLeuCyLysValHisIleProSerHisIleAspPheArgAspCys 713
 Db 1452 CATGTCTGGAGAGGCTGTGAAACACCAACCAACCGCGACTTCAGGGAGCTGC 1511
 QY 714 ArgAlaAlaLeuValagLulIleThrGluMetValIleGlnLeuHisIleGlyThrMetIleLys 733
 Db 1512 AGAGCTCGGCTGGCGAGATACAGAGATGGTGGCACAGCTGACATGATATCAAG 1571
 QY 734 MetGluAsnPheGlnLysLeuHisIleGlnLeuLysLysAspLeuIleGlyIleAspAsnLeu 753
 Db 1572 ATGGAGAACTTCCGAAGCTGCATAGCTCAAGAAAGTCTGATCGCATTAACATCTT 1631
 QY 754 ValValProGluArgGluPheIleArgLeuGlySerLeuSerLysLeuSerGlyLysGly 773
 Db 1632 GTGACCCCAAGGAAGGAGTTATCCGCTGGCGAGCTTCAGCAAGCTTCGCGGAAGGG 1691
 QY 774 LeuGlnGlnArgMetPhePheLeuPheAsnAspValLeuLeuTyThrSerArgGlyLeu 793
 Db 1692 CTTACAGACGCGCATGTTCTTCTCTTCAACATGTTGTGTATACAGCCGCGGAGCTG 1751
 QY 794 ThrAlaSerLeuGlnIlePheLysValHisIleGlnLeuProLeuTyThrLysMetThrIleGlu 813
 Db 1752 ACCGACATTAATCAGTTAAAGTCCAGGACAGCTCCACATTAAGCATGATGATGAG 1811
 QY 814 GluSerGluAspGluTyrGlyValIleProHisCysLeuThrLeuAspGlyGlnArgGlnSer 833
 Db 1812 GAGAGTAGAGAGAGGTGGGTGTGCCCCACTGCTGACCTCGCGGGGCGACGGCAATCT 1871
 QY 834 IleIleValAlaAlaSerSerArgSerGluMetGluLysTyrValIleAspIleGlnMet 853
 Db 1872 ATCATCGTGGCTGCCAGTTCGCCGTCAAGATGAGAAATGATGAGAGGACATCCAGATG 1931
 QY 854 AlaIleAsnLeuAlaGluLysSerSerSerProAlaProGluPheLeuAlaSerSerPro 873
 Db 1932 GCCATTGACCTGGAGAGAAAGCAATGAGACCCCGCGAGCTCTTGGCAGAGAGCCC 1991
 QY 874 ProAspAsnLysSerProAspGluAlaThrAlaAlaAspGlnIleSerGluAspAspLeu 893
 Db 1992 CTTGACAAACAGTCCCGGATGAAAGCCACAGCGGCTGACAGAGTCAAGAGATGACTG 2051
 QY 894 SerAlaSerArgThrSerLeuGluArgGlnAlaProHisArgGlyAsnThrMetValHis 913
 Db 2052 AGTGCCTCTCGAGACATCGCTGAGACGCCAGGCCCTCAGCGTGCAACAGATGCTGCAC 2111
 QY 914 ValCysThrPheIleArgAsnThrSerValIleSerMetValAspPheSerIleAlaValGluAsn 933
 Db 2112 GTGTGCTGGCATTCGACACACCAAGTGTCTCCATGGTGAATTGACATTCGCTGGAAAC 2171
 QY 934 GlnLeuSerGlyAsnLeuLeuArgLysPheLysAsnSerAsnGlyTyrTrpGlnLysLeuTyr 953
 Db 2172 CAGCTCTCTGGGAATCTGCTGAGAAAGTTCAAGAACAGCAATGCTGGCAGAACTGTGG 2231
 QY 954 ValValPheThrAsnPheCysLeuPhePheTyrlLysSerHisIleAspAsnHisIleProLeu 973
 Db 2232 GTGGTATTACCAAACTTCTGCGCTGTCTTCAAAATCACACGAGACATGACCCCTT 2291
 QY 974 AlaSerLeuProLeuLeuGlyTyrlSerLeuThrIleProSerGlyLysGluAsnIleGln 993
 Db 2292 GCCAGCGCTGCTGGCTGGGTACTCTCTCAACATCCCTCTGAGTGAAGAAATCATCAC 2351
 QY 994 LysAspTyrlValPheLysLeuHisIlePheLysSerHisIleValTyrlTyrlPheArgAlaGlnSer 1013
 Db 2352 AAGGACATATGTTCAGAGCTGCAATTCAGTCCACAGGTATCTTCAAGGCTGAAAGC 2411
 QY 1014 GluTyrlThrPheGluArgTyrPheMetGluValIleArgSerAlaThrSerSerAlaSerArg 1033
 Db 2412 GAATACACATTTGAAGGTGATGGAAGTATCCGAAGTGCACACAGCTCAGCTCTCGG 2471

QY 1034 ProHisValLeuSerHisLysGluSerLeuValTyrl 1045
 Db 2472 GCCCAGATCTTGAAGTCAAAAGAAATCTACCTGTAC 2507
 RESULT 9
 LOCUS CO895396
 DEFINITION Sequence 16 from Patent EP1471153.
 ACCESSION CO895396
 VERSION CO895396.1 GI:55467577
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homnidae; Homo.
 REFERENCE
 1
 AUTHORS Certa, U., Foser, S. and Weyer, K.
 TITLE Transcriptional activity assay
 JOURNAL Patent: EP 1471153-A 16 27-OCT-2004;
 F. HOFMANN-LA ROCHE AG (CH)
 FEATURES
 source
 1. 3997
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 ORIGIN
 Alignment Scores:
 Pred. No.: 1,42e-231 Length: 3997
 Score: 2941.50 Matches: 580
 Percent Similarity: 70.0% Conservative: 162
 Best Local Similarity: 54.7% Mismatches: 261
 Query Match: 53.8% Indels: 57
 Gaps: 2
 DB: 14
 US-09-555-342B-2 (1-1045) x CO895396 (1-3997)
 QY 1 MetGlyGluIleGluGlnArg-----ProThrProGlySerAlaGluValAla 16
 Db 118 ATGGGGAGATAGAGAGAACATACAGAGTCTTCGACAGCTCAGAGGATGCGTTGGGTGCC 177
 QY 17 ProGluAsnSerGlyIleSerThrLeuGluArgGlyGlnLysProProThrProSer 36
 Db 178 CAGACCCCTGTGGAGATGAGACCTTGAGCTGGGAGACTCTCTTGGCCAGATGCAA 237
 QY 37 GlyLysLeuValSerIleLysIleGlnMetLeuAspAspThrGlnGluAlaPheGluVal 56
 Db 238 GAGAAAGACCTGACACCTCAGAGTAAAGCTGCTGGAACAACATGAAATATTTGACATT 297
 QY 57 ProGlnArgAlaProGlyLysValLeuLeuAspAlaValCysAsnHisIleAsnLeuVal 76
 Db 298 GAGCTTAAATGCGATGGCCAGATATTACTGACACAGTGTGGAAGCGTTTAAACCTGTGA 357
 QY 77 GluGlyAspTyrlPheGlyLysGluPheProAspHisLysLysIleThrValTyrProLeuAsp 96
 Db 358 GAATGATACATCTTCGGAGATGAGATTCAAAATATCTAGTCTTACTGATTTGGCTTGA 417
 QY 97 LeuLeuLysProIleValLysGlnIleArgArgProLysHisIleValLysPheVal 116
 Db 418 CCGATGAAACCATCATTAAGCAAAATAGCAAGGCCAAAGATGTGTGCTCGCTAGCT 477
 QY 117 ValLysPhePheProProAspHisIleThrGlnLeuGlnGluGluLeuThrArgTyrlLeuPhe 136
 Db 478 GTAAATTTTTCACCTGTGATCTGTGACCTTCMAAAGAAATATACAAAGATATCTTGT 537
 QY 137 AlaLeuGlnValLysGlnAspLeuAlaGlnGlyArgLeuThrCysAsnAspThrSerAla 156
 Db 538 GCTTGGACATTAAGAGAGACTGCTGGAAGAGCTTTGACCTGTGTGACACACAGCG 597
 QY 157 AlaLeuLeuIleSerHisIleValGlnSerGluIleGlyAspPheAspGluAlaLeuAsp 176
 Db 598 GCCCTTCTACAGTCCCATCTCTGCAAGTCCGAAATAGAGATATGATGAAACGCTGAC 657

QY 177 ArgGluHisLeuAlaLysAsnLysTyrIleProGlnGlnAspAlaLeuGluAspLysIle 196
 DB 658 CGAGAGCACTTCAAGATGAAAGATATTCCTGGCCAGCAGCACTGCTTGAAGATTA 717
 QY 197 ValGlnPheHisValAsnHisIleGlyGlnThrProAlaGlnSerAspPheGlnLeu 216
 DB 718 CTGAATTCATCAGAAAGCAGTGGGCGAAGACCTGCTGATGCGATTTCCAGGTCTC 777
 QY 217 GlnIleAlaArgArgLeuGlnMetTyrGlyIleArgLeuHisProAlaLysAspArgGlu 236
 DB 778 GAATTCCTCGAAATGTGAATGTACGCGATCAGATTTCACTGCTTGACAGGAA 837
 QY 237 GlyThrLysIleAsnLeuAlaValAlaAsnThrGlyIleLeuValPheGlnGlyPheThr 256
 DB 838 GGAACCAAGATTCACACTGGCACTTCCACATGGGTCTACTCGTTCCTCCAGGCAACACC 897
 QY 257 LysIleAsnAlaPheAsnTrpAlaLysValArgLysLeuSerPheLysValArgPhe 276
 DB 898 AAATCAACACTTTCACCTGATCGTCCAAAGTCCGTAAACTTAAGCTTCAAGAGAAAGATT 957
 QY 277 LeuIleLysLeuArgProAspAlaAsnSerAlaTyrGlnAspThrLeuGlnPheLeuMet 296
 DB 958 CTATCAACTTCATCCAGAGGTTCATGACCTTACCAAGACACATTAAGAAATTTTGTG 1017
 QY 297 AlaSerArgAspPheCysLysSerPheTrpLysIleCysValGlnHisValaPhePhe 316
 DB 1018 GGTAGTAGAATGAAATGAAACTTCTGAAGATTTGTGTGAGATTCACACTTTT 1077
 QY 317 ArgLeuPheGlnGlnProLysProLysProLysProValLeuPheSerArgLysSer 336
 DB 1078 AGACTTTGGACCACTTAAGCCAAAGCAAAAGCCGCTTCTTCCAGCCGGGCTCCTCC 1137
 QY 337 PheArgPheSerGlyArgThrGlnLysGlnValLeuAspTyrValLysGlnGlyHis 356
 DB 1138 TTGAGATACAGTGGAAAGACTCAGAAACAATGATTAATTTCAAAGCAGTGAATG 1197
 QY 357 LysLeuValGlnPheGlnArgLysHisSerLysIleHis---SerIleArgSerLeu--- 374
 DB 1198 AAGAGATTCATATGAAAGAAAGAGCAGCAAGACCCACAGCTCCGTCCAGCTGACT 1257
 QY 375 AlaSerGlnProThrGlnLeuAsnSerGlnValLeuGlnGlnSerGlnInSerThrSer 394
 DB 1258 GCAGACCTACCAAAACAG-----AGC 1278
 QY 395 LeuThrPheGlyGlnGlyAlaGlnSerProGlyGlyGlnSerCysArgArg----- 411
 DB 1279 ATCTCATTTCCCGAGGATTTGAGGACTGCTGCCATCCCATCTTCAAGGAATGCTTTTAC 1338
 QY 412 -----GlyLysGlnProLysValSerAlaGlyGlnProGly----- 423
 DB 1339 TCGCTCTCTCCCTCCACTCTGCTCCCTCTGCTGCTGCAAGTTTAAAGACAGCAGC 1398
 QY 424 -----SerHisProSerProAlaProArgArgSerProAlaGlnLysGlnAlaAsp 441
 DB 1399 TCCCTACACATCCCGAGTTCCTTACGTCAAGAGTCCAGCTCAGAGAGCCCAAGTGA 1458
 QY 442 GlyAlaAlaSerAlaProThrGlnGlnGlnGlnGlnValValLysAspArgThrGlnGln 461
 DB 1459 GCAGTGGCTGGAGGCCCGAC-----ACACCATCG 1488
 QY 462 SerLysProGlnProPro-----GlnProSerThrGlySerLeuThrGlySerPro 478
 DB 1489 GCCACAGCCCTGGGGCCCCCGCAGCTCCAGCTGCTGTCAGGCTTTCCAGAAAGTCC 1548
 QY 479 HisLeuSerGlnLeuSerValaAsnSerGlnGlyValAlaProAlaAsnValThrLeu 498
 DB 1549 CAGCTTCTCCCTCCACCGGAGAGGCCCTGAGTGTGAGCCCTGCA-----TTT 1599
 QY 499 SerProAsnLeuSerProAspThrLysGlnAlaSerProLeuLysSerProLeuLeuAsn 518
 DB 1600 CAGGTGCTTTGGGCGCAGCTGAACAGGCTCATCCCTCCACTCTGAGCCCTGTCTCAGT 1659

QY 519 AspGlnAlaCysProArgThrAspArgGluAspGlnGlyValArgArgLysArgPheProThr 538
 DB 1660 GATGCTGGGGGAGCCCGGATGACTCGAG---GAGCCAGAGACAAAGGCGTGGCTGCA 1716
 QY 539 AspLysAlaTyrPheIleAlaLysGlnValSerThrThrGlnArgThrTyrLeuLysAsp 558
 DB 1717 GACGAGGCTTACTTCTACATGCAAGAGATTCGCTGCTACAGAAAGCAATCTCCCAAGAT 1776
 QY 559 LeuGlnValIlePheSerTrpPheGlnSerThrValSerLysGlnAspAlaMetProGlu 578
 DB 1777 TTGAAGTTATTAACGTGTGTTCCGAGCGCAGTGTGAAGAGAGCCGCACTTGGC 1836
 QY 579 AlaLeuLysSerLeuIlePheProAsnPheGlnProLeuHisLysPheHisThrAsnPhe 598
 DB 1837 ACTGTATAGCGGTGCTTCTTCCATCATGATCCCATGTATAGTTTCCACAGAGCTTC 1896
 QY 599 LeuLysGlnIleGlnGlnArgLeuAlaLeuTrpGlnGlyArgSerAsnAlaGlnIleArg 618
 DB 1897 CTGCGGAGGTGGAGCAGAGGCTGCGCATCTGGAAAGGCCCTCCAAAGCCCAACAAAA 1956
 QY 619 ---AspTyrGlnArgIleGlyAspValMetLeuLysAsnIleGlnGlyMetLysHisLeu 637
 DB 1957 GCGATGATCAAGAAATCGGGACATCTGCTCAGAAACATGCGCCATTAAGAGCTTT 2016
 QY 638 AlaAlaHisLeuTrpLysHisSerGlnAlaLeuGlnAlaLeuGlnLysLysSer 657
 DB 2017 ACCAGCTACTTCCAAAGACATGACGAGTCTTACAGAACTGGAAAGGCTACCAAGCGC 2076
 QY 658 SerArgArgLeuGlnLysAsnPheCysValArgPheGlnGlnLysValCysTyrLeuPro 677
 DB 2077 TGTAAAGTTGGAGGACAGTGTACAGAGATTGACCTCAGAAAGTCTGCTACTTGGCT 2136
 QY 678 LeuAsnThrPheLeuLeuArgProLeuHisValGlyLeuMetHisTyrTyrGlnValLeuGlu 697
 DB 2137 CTCAACACGTTCTCTGTGAAGCCATCCAGCGGCTCTGTGACTACCGCTGCTGCGC 2196
 QY 698 ArgLeuCysLysHisValProProSerHisValaAspPheArgArgValaAlaLeu 717
 DB 2197 CGCTATGGGAACTATTAAGCCCCGGGCAACATGACTAGCTAGCCATGACGCCCTG 2256
 QY 718 AlaGlnIleThrGlnMetCysValaAlaGlnLeuHisGlyThrMetIleLysMetGluAsnPhe 737
 DB 2257 AAGCCATCAACAGAGGTGACCAACACATCAAGCACATTTCTCATCCGGCTGAGAACTTG 2316
 QY 738 GlnLysLeuHisGlyLeuLysLysAspLeuIleGlyIleAspAsnLeuValValProGly 757
 DB 2317 CAGAGCTTACCGAGGCTGCAAGCGGACCTGGTGGGATGAGAACTCATGCTCTGCGC 2376
 QY 758 ArgGlnPheIleArgLeuGlySerLeuSerLysLeuSerGlyLysGlyLeuGlnGlnArg 777
 DB 2377 AGGGAATTCATCCGTGAGGGCTGCTTCAAGCTCACCAAGAAAGGCGCTGACAGAGG 2436
 QY 778 MetPhePheLeuPheAsnAspValLeuLeuTyrThrSerArgGlyLeuThrAlaSerAsn 797
 DB 2437 ATGTTTTTCTGTTCACATATGTGTCGTACACAAAGAGAAAGTTCAGAGGACCAAGC 2496
 QY 798 GlnPheLysValHisGlyGlnLeuProLeuTyrGlyMetThrIleGlnGlnSerGlnLys 817
 DB 2497 CACTTCGGAATCGGGGCTCTCTCCCTCCAAAGCATCTGTGTGAAGAAATGATTAAC 2556
 QY 818 GlnTrpGlyValProHisCysLeuThrLeuArgGlyGlnArgGlnSerIleIleValAla 837
 DB 2557 GAGTGTCTGTTCCACATGTTTCAACATCTACCGCGCTCAGAAACAAATCGTGTGCA 2616
 QY 838 AlaSerSerArgSerGlnMetGlyLysTrpValGlnAspIleGlnMetAlaIleAspLeu 857
 DB 2617 GCCAGCATCTCGGTGAGAAAGAGTGAAGTCTGAGACTGAACCTCCGGAATCCAAAGCA 2676
 QY 858 AlaGlnLysSerSerSerProAlaProGluPhe-----LeuAlaSerSerPro 873
 DB 2677 GCCAAAGTGGCGGTACACGCGCCCTGCACTGCCAGCGGCACTGTGTGCACTGCTCC 2736
 QY 874 ProAspAsnLysSerProAspGlnAlaThrAlaAlaAspGlnGlnLysArgLysAspLeu 893


```
Db 2737 CCC-----AGATCCCCCAAGAG---GTATCTCGGAGCGAGGTCAAGAAATGATGCT 2787
Qy 894 SerAlaSerArgThrSerLeuGluArgGlnAlaProHlaArgGlyAlaThrMetValHis 913
Db 2788 CCGGGGTCTCCGACCTCTCTGGAGGGGATGCCAGCCAGCCGGGCCAACACCAATGAC 2847
Qy 914 ValCysThrPheAspGlnThrSerValSerMetValaPheSerIleAlaValGlnAsn 933
Db 2848 GTGTCTGGTACCCGAAACCAAGCGGTCTCGAGGACAGCCAGTGCAGCTGTGAGAAC 2907
Qy 934 GlnLeuSerGlyAsnLeuLeuArgGlyPheLeuAsnSerAsnGlyTyrGlnLeuLeuTyr 953
Db 2908 CAGCTTCAGAGTATCTGCTAAAGAAAGTTCAAAACAGTACGTGCTGGCAAACTCTGG 2967
Qy 954 ValValaPheThrAsnPheCysLeuPhePheTyrIleSerHisGlnAspAsnHisProLeu 973
Db 2968 GTCTCTTTTCCAACTTCTGTCTTCTTCTTCAAAACTCATCGAGTACTACCACTG 3027
Qy 974 AlaSerLeuProLeuGluGlyTyrSerLeuThrIleProSerGluSerGluAsnIleGln 993
Db 3028 GCCAGCTCCCGCTGCTGGCTAGCGGTGAGCAATCCCAAGGAGCCGATGCGCATACAC 3087
Qy 994 LysAspTyrValPheLeuLeuHisPheLeuSerHisValTyrTyrPheArgAlaGluSer 1013
Db 3088 AAGACATATGTTTCAAGCTCAGATTCAATCCCAAGTCTACTCTTCCGGCTGAGAGC 3147
Qy 1014 GluTyrThrPheGluArgTyrMetGluValIleArgSerAlaThrSerSerAlaSerArg 1033
Db 3148 AAGTACACATTTGAAGGTGATGAGGTGATCCAGGAGGCCAGACACTCAGCCGGAGG 3207

RESULT 10
AB018336 3997 bp mRNA linear PRI 10-JAN-2004
LOCUS Homo sapiens mRNA for KIAA0793 protein, partial cds.
ACCESSION AB018336
VERSION AB018336.1 GI:3882306
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Nagase,T., Ishikawa,K., Suyama,M., Kikuno,R., Miyajima,N.,
Taniike,A., Kotani,H., Nomura,N. and Ohnara,O.
TITLE Prediction of the coding sequences of unidentified human genes. XI.
The complete sequences of 100 new cDNA clones from brain which code
for large proteins in vitro
JOURNAL DNA Res. 5 (5), 277-286 (1998)
PUBMED 9872452
REFERENCE 2 (bases 1 to 3997)
AUTHORS Ohnara,O., Suyama,M., Nagase,T., Ishikawa,K. and Kikuno,R.
TITLE Direct Submision
JOURNAL Submitted (08-OCT-1998) Otsu Ohara, Kazuo DNA Research Institute,
Laboratory of DNA Technology; Yana 1532-3, Kizakazu, Chiba
222-0812, Japan (E-mail:cdna@infokazuo.or.jp, Tel: +81-438-52-3913,
Fax: +81-438-52-3914)
FEATURES
source
1.3997
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="hk05692"
/sex="male"
/tissue_type="brain"
/clone_id="pbluescriptII SK plus"
/dev_Stage="adult"
1..3997
/gene="KIAA0793"
<115..3282
/gene="KIAA0793"
/note="Start codon is not identified"
```

```
/codon_start=1
/product="KIAA0793 protein"
/protein_id="BA34513.2"
/db_xref="GI:40788367"
/translation="RMGEICGYRVLQTAQMRILGAQTPGVSVLSEPGQTLLPRMQEKH
LHLRVKLDNTWMEIFDIEPKDGDVLTQWVKRLNLTVECDYFGMEFQNTQSYIMLEP
MKPIRQIRRRKVVNLAVKFFPPDQGLQOEYTRYLFAQLKRLDLERLTCADTT
AALLTSLHLSOSIGDYDETDREHLKXNEYLPGQOGLLEKILEPHQGVOTPAESP
QVLEIAKMEYGRFHMASDRGCTQLASVHNGVLPQGTIKIMFNMSKRYKLSF
KRRFLIKLPEVHGPYQDTLEBELGLSDRECKRPFWKCVETHTFFRLDDPKRPAKV
FSGSGSFYSGRQKQVDFKDSQMKRLPYERRSHKTHSVATLADLPKQISFP
EGLRTPASPSANAFYLSFSTLVPSELPFCKSSSLTDPQVSVYVSPAAERNSGV
AGSPDTPAQPLQPGSLTSPQSPSSRSKPLSLSPAFVPLPAEGSSAP
LSPVLSADGAGKDCSEPRHKVPADNAVFIYEKILATERTYLKLELVTPRSAY
VKEADAMPATMTLLFSNIDPIYERHNGRLFEVQRLALMRGSPKATNGHRTGIDL
LRMRQLKERTSYFORHDEVLTLEKATKCKLEAYKREBLQKVTLYPLNTLPLP
IQRLNHYRLRLRLCGHYSQGNHDVADCHDLAKATEVTTLOHILRLNLOQLTEL
QRDLVGENLEIARGREFIRGCILKCLKLQRMFPLFSDMLLYTSKGVAGSHFRI
RGLLPLQGMVSESDNEMSVPHCFETIYAAQTIYVAASTLEKEMKMLDINSALQAAK
SGDDTPALPGRITVCTRPPRSFNVLSLESDPDARGVRSLSLEGHQRANTMHWCV
YRNTSVSRADHSAAVENQLSGYLLRKRKSHGQKLMVFTNRLPFYKTHODDYPILA
SLPLGYSVSIPEADGITHQVYFKLQFKSHVYFFRAESKTYTERMMEVIGASSSAG
RASPVDGQPPSSGLEGMVRKEE"
ORIGIN
Alignment Scores:
Pred. No.: 1,42e-231 Length: 3997
Score: 2941.50 Matches: 580
Percent Similarity: 70.0% Conserved: 162
Best Local Similarity: 54.7% Mismatches: 261
Query Match: 53.8% Indels: 57
DB: Gaps: 14
US-09-555-342B-2 (1-1045) x AB018336 (1-3997)
Qy 1 MetGlyGluIleGluGlnArg-----ProThrProGlySerArgLeuGlyAla 16
Db 118 ATGGGGAGATAGAGAGACATACAGAGTCTCTGAGACTCGAGGAGATCGCTTGAGGCC 177
Qy 17 ProGluAsnSerGlyIleSerThrLeuGluArgGlyGlnIlyProProProThrProSer 36
Db 178 CAGACCCCTGTGGAGATTACACCCCTTGAGCTTGGCGAGACTCTCTTCCCGCAGAAATCAA 237
Qy 37 GlyIysLeuValSerIleLeuIleGlnMetLeuAspAspThrGlnGluAlaPheGluVal 56
Db 238 GAGAAAGCACTGCACTCAAGATTAAGCTGTGCAACAACCATGAGAAATATTGACATT 297
Qy 57 ProGlnArgAlaProGlyIlyValLeuLeuAspAlaValCysAsnHisLeuAsnLeuVal 76
Db 298 GAGCTTAATGCCATGGCCAGGTATTACTGACACAACTGTGGAAGCGTTTAACCTGTGTA 357
Qy 77 GluGlyAspTyrPheGlyLeuGluPheProAspHisIlyIlyIleThrValTyrLeuAsp 96
Db 358 GAATGTACTACTTCCGAGATGAGATTCAAAATCTCAGCTCACTAGATTTGGCTTGAA 417
Qy 97 LeuLeuAspProIleValIlyGlnIleArgArgProIlyHisValValIlyAspPheVal 116
Db 418 CTTATGAACCCATCATTAAGCAAAATGCAAGGCAAAAGATGGTGGCTTCGCTTGCT 477
Qy 117 ValIysPhePheProProAspHisIleThrGlnLeuGlnGluIleuThrArgTyrLeuPhe 136
Db 478 GTAAATATTTTTCACACTGATCCCTGTCAGCTACAAAGAAATATACAGATCTTCTTT 537
Qy 137 AlaLeuGlnValIlyGlnAspLeuAlaGlnGlyArgLeuThrCysAsnAspThrSerAla 156
Db 538 GCCCTTGCACTTAAGAGAGACCTGCTGAGAGGTTTGACCTGTGTCAGCACACGCG 597
Qy 157 AlaLeuLeuIleSerHisIleValGlnSerGluIleGlyAspPheAspGluAlaLeuAsp 176
Db 598 GCCCTTCTCAGCTCCATCTCTGCGACGTGGAATTAAGAAATTACGTGAACCGCTGGAC 657
Qy 177 ArgGluHisLeuAlaIlyAsnIlyTyrIleProGlnGlnAspAlaLeuGluAspIlyIle 196
```

Dh 658 CGAGAGCACTCAAAAGTAAGAGATATTGCTGGCCAGCAGCACTGCTTGAAGAATA 717
Qy 197 ValGluPheHsiBsaHsiSileGlyGlnThrProAlaGluSerAerPheGlnLeuLeu 216
Db 718 CTGAATTCATCATGAAAGCAGTGGCCAGACCTGCTGATCGATTCAGGTTCAGGTCTC 777
Qy 217 GlnIleAlaArgAerGluMetTyrGlyIleArgLeuHsiProAlaLysAerGlu 236
Db 778 GAATTCCTGAAAGATTGAAATGTACGGCATCATGATTCAATGCTTGCACAGGAA 837
Qy 237 GlyThrLysIleAsnLeuAlaValAlaAsnThrGlyIleLeuValPheGlnGlyPheThr 256
Db 838 GGAACCAAGATTCAATCGCATGTTCCACATGGTGTACTCGTGTCCAGGCGACACC 897
Qy 257 LysIleAsnAlaPheAsnThrAlaLysValArgLysLeuSerPheLysAerGlyAerPhe 276
Db 898 AAAATCAACCTTCAACTGATCGTCCAGGTCCTGTAACCTTCAAGCTTCAAGAGAAAAGATT 957
Qy 277 LeuIleLysLeuArgProAerAlaAsnSerAlaTyrGlnAerThrLeuGluPheLeuMet 296
Db 958 CTATCAAACTTCATCCAGAGCTTCATGACCTTACAGACATTAAGATTTTCTTG 1017
Qy 297 AlaSerArgAerPheCysLysSerPheTyrLysIleCysValGlnHsiBsaAerPhe 316
Db 1018 GGTAGTAGATGATGATGTAAGAACTTCTGGAAGATTGTGTGGAGTATCACACCTTTT 1077
Qy 317 ArgLeuPheGluGlnProLysProLysProLysProAlaLeuPheSerArgLysSer 336
Db 1078 AGACTTTGACCAACTTAAGCCAAAGCAAAAGCCCTTCTTCAACCGGAGCTCTCC 1137
Qy 337 PheArgPheSerGlyArgThrGlnLysGlnValLeuAerTyrValLysGlnGlyLys 356
Db 1138 TTCAAGATACAGTGGAAACATCAGAAACAATACTAATTATTTCAAGACAGTGAATG 1197
Qy 357 LysLysValGlnPheGluArgLysHsiSerLysIleHsi---SerIleArgSerLeu--- 374
Db 1198 AAGAAGATTCATATGTAAGAAGGACAGACAAAGCCACAGCTCGCTCAGCTGACT 1257
Qy 375 AlaSerGlnProThrGluLeuAsnSerGluValLeuGlnGlnSerGlnInsThrSer 394
Db 1258 GCAAGCTTACCAAAACAG-----AGC 1278
Qy 395 LeuThrPheGlyGlnGlyAlaGluSerProGlyGlyGlnSerCysAerArg--- 411
Db 1279 ATCTCATTCCTCCAGAGGATTGAGGACTCGCTCCCTCCCATCTTCAAGGAATGCTTTAC 1338
Qy 412 -----GlyLysGlnProLysValSerAlaGlyGlnProGly----- 423
Db 1339 TCGCTCTCTCCCTCACTGCTGCTCCCTGCTGCTGCGCAGATTTAAGACAGACAGC 1398
Qy 424 -----SerHsiProSerProAlaProArgAerSerProAlaGlnLysGlnAlaAer 441
Db 1399 TCCCTCAACAATCCCAAGGTTTCTTACGTCAAGAGTCAGCTGACAGAGGCGAGTGA 1458
Qy 442 GlyAlaAlaSerAlaProThrGlnGlnGlnGlnGlnValValLysAerArgThrGlnGln 461
Db 1459 GCAAGTGGCTGAGGCCCGAC-----AACCATCG 1488
Qy 462 SerLysProGlnProPro-----GlnProSerThrGlySerLeuThrGlySerPro 478
Db 1489 GCCAGCCCTCGGAGCCCGCCGCACTCCAGCTGGTCCAGGCTTTCACAGAAAGTCC 1548
Qy 479 HsiLeuSerGluLeuSerValaAsnSerGlnGlyGlyValAlaProAlaAsnValThrLeu 498
Db 1549 CAGCTTCTCTCCACGCGAAGAGCCCTGAGTGAAGCTTGA-----TTT 1599
Qy 499 SerProAsnLeuSerProAerThrLysGlnAlaSerProLeuLysSerProLeuAsn 518
Db 1600 CAGGTGCTTTGGCCAGCGAAGAGGCTCATCCCATCTCGAAGCCCTGCTCAGT 1659
Qy 519 ArgGlnAlaCysProArgThrAerAerPheGlnAerGlnGlyArgLysAerPheProThr 538
Db 1660 GATGCTGGCGAGCGGAGTGAAGTCTGAG---GAGCCAGACACAAAGCGCTGCTGCA 1716

Qy 539 AspLysAlaTyrPheIleAlaLysGlnValSerThrThrGlnArgThrTyrLeuLysAer 558
Db 1717 GACAGAGCTTACTTCAATGTCAAAGAGATTCTGCTACAGAACCAATCTCCAGGAT 1776
Qy 559 LeuGlnValIleThrSerTyrPheGlnSerThrValSerLysGlnAerAlaMetProGlu 578
Db 1777 TTGAAGATTATTAACGTGTGGTTCCGACGGCAGTGTGAAGAGAGCGCATGCTGCG 1836
Qy 579 AlaLeuLysSerLeuIlePheProAsnPheGluProLeuHsiBysPheHsiThrAerPhe 598
Db 1837 ACTGTATGACGTGTCTTCTCAACATGATCCATCATATGATGATTCACAGAGCTTC 1896
Qy 599 LeuLysGlnIleGlnGlnArgLeuAlaLeuTyrGlnGlyLysSerAsnAlaGlnIleArg 618
Db 1897 CTGCGGAGGTGAGCAGAGGCTGGCATCTGGGAAGGCGCTCCAAAGCCACACAAA 1956
Qy 619 ---AspTyrGlnArgIleGlyAerValMetLeuLysAsnIleGlnGlyMetLysHsiLeu 637
Db 1957 GGCAGTCATCAAGATCGGAGCATCTGCTCAGGAACATGCGCCAGTTAAAGAGTTT 2016
Qy 638 AlaAlaHsiLeuThrLysHsiSerGlnAlaLeuGlnAlaLeuGlnLysGlnLysSer 657
Db 2017 ACCAGCTACTTCCAAAGACATGACGAGTCTTAACAGAACTGGAAGGCTACCAACGC 2076
Qy 658 SerArgArgLeuGlnLysPheCysAerAerPheGlnLeuGlnLysValCysTyrLeuPro 677
Db 2077 TGTAAAGTTGAGGACAGCATGTAACAAGATTGAGCTGCAAGAGCTCTCTGCT 2136
Qy 678 LeuAsnThrPheLeuLeuArgProLeuHsiBysLeuMetHsiTyrLysGlnValLeuGln 697
Db 2137 CTCAAACGTTCTGCTGTAAGCCCATCCAGCGGCTGTGACATACCGCCCTGCTGCGC 2196
Qy 698 ArgLeuCysLysHsiBsiProProSerHsiAlaAerPheArgAerPheArgAlaAlaLeu 717
Db 2197 CGCTATGCGGACATTACAGCCCGGACCATGATACGCTGATGCTCATGACGCTG 2256
Qy 718 AlaGlnIleThrGlnMetValAlaGlnLeuHsiGlyThrMetIleLysMetGlnAsnPhe 737
Db 2257 AAGGCATCAGAGGTTGACACCAACATACAGCAATCTCATCCGCGTGAAGACTG 2316
Qy 738 GlnLysLeuHsiGlnLeuLysLysAerLeuIleGlyIleAsnAsnLeuValValProGly 757
Db 2317 CAGAACTTAACGAGGCTGACCGGAGCTGTGGCATGAGAACTCATTCCTCTGCGC 2376
Qy 758 ArgGlnPheIleArgLeuGlnSerLysLeuSerGlyLysGlyLeuGlnGlnArg 777
Db 2377 AGGAGTTCATCTCGTAGGCTGCTTCAACAGCTCACAAAGAGGCTGCGACAGAGG 2436
Qy 778 MetPhePheLeuPheAsnAerValLeuLeuTyrThrSerArgLysLeuThrAlaSerAsn 797
Db 2437 ATGTTTTCCTGTTCCAGATATGTTGCTGTACAAAGAAAGAGTGCAGGAGCACAGC 2496
Qy 798 GlnPheLysValHsiGlyGlnLeuProLeuTyrGlyMetThrIleGlnGlnSerGluAer 817
Db 2497 CACTTCGAGTCGGGCGCTCTTCCCTTCCCAAGGATGCTGTGGAAGAAATGATTAAC 2556
Qy 818 GlnTyrGlyValProHsiCysLeuThrLeuAerGlyGlnArgGlnSerIleIleValAla 837
Db 2557 GAGTGTCTGTTTCAACATGTTTCACTTACATTCACGCGCTCAAGAAACAATGTGTGGCA 2616
Qy 838 AlaSerSerArgSerGlnMetGlnLysTyrValGlnLysIleGlnMetAlaIleAerLeu 857
Db 2617 GCCAGCATCTCGTGAAGAAAGAGTGAATCTGGAACCTGAACTCCGAGATCCAAAGCA 2676
Qy 858 AlaGlnLysSerSerProAlaProGluPhe-----LeuAlaSerSerPro 873
Db 2677 GCCAAGAGTGCAGTACACGCGCTTGCATCCAGCGCGCATCTGTGCTACCTGCTCC 2736
Qy 874 ProAerAerLysSerProAerGlnAlaThrAlaAlaAerGlnGlnLysGlnAerAerLeu 893
Db 2737 CCC-----AGATCCCAACAG---GATCTCTGGAAGAGGATGTAAGTGAATGCT 2787

Qy 894 SerAlaSerArgThrSerLeuGluArgGlnAlaProHisArgGlyAsnThrMetValHis 913
Db 2788 CGGGGTGTCGGACCTCCCTGGAAGGCGATGCGACACCGGGCCAAACACCAATGAC 2847
Qy 914 ValCysTrpHisArgAsnThrSerValSerMetValAspSerIleAlaValGluAsn 933
Db 2848 GTGGTGTGTAACCGAACCAGCGCTGCCAGGGCAGACCAAGTGCAGCTGTGAGAAC 2907
Qy 934 GlnLeuSerGlyAsnLeuLeuAlaGlyPheLeuAsnSerAsnGlyTrpGlnLeuLeuTrp 953
Db 2908 CAGCTTTCAGAGATCTGCTTAAGAAAGTTCMAAAACGTACTGCTGGCAGAACCTCTGG 2967
Qy 954 ValValPheThrAsnPheCysLeuPhePheTyrLeuSerHisGlnAspAsnHisProLeu 973
Db 2968 GTGCTCTTACCAACTTCTGTTTCTTCTTACAAACTCATCTCAGATGACTCCACTG 3027
Qy 974 AlaSerLeuProLeuLeuGlyTyrSerLeuThrIleProSerGluSerGluAsnIleGln 993
Db 3028 GCCAGCTCCCGCTGGGCTACAGCGTACAGCGTACCCAGGAGGCGCATGAC 3087
Qy 994 LysAspTyrValPheLeuLeuHisPheLeuSerHisValTyrTyrPheArgAlaGluSer 1013
Db 3088 AAGACTATGTGTTTCAAGCTCCAGTCCAAATCCACGCTACTTCTCCGGCTGAGAGC 3147
Qy 1014 GluTyrThrPheGluArgTyrTrpMetGluValIleArgSerAlaThrSerSerAlaSerArg 1033
Db 3148 AAGTACATTTGAAAGGTGATGAGATGATCCAGGGGCGCAGACGCTGACCGGAGG 3207
RESULT 11
LOCUS CQ725044 3997 bp DNA linear PAT 03-FEB-2004
DEFINITION Sequence 10978 from Patent WO2068579.
ACCESSION CQ725044
VERSION CQ725044.1 GI:42285901
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.
REFERENCE
AUTHORS Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
TITLE Ktfs, such as nucleic acid arrays, comprising a majority of
humanexons or transcripts, for detecting expression and other uses
thereof
JOURNAL Patent: WO 02068579-A 10978 06-SEP-2002;
FEATRES location (NY) (US)
source 1. 3997
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
pred. No.: 3 66e-231 Length: 3997
Score: 2936.50 Matches: 580
Percent Similarity: 69.9% Conservative: 161
Best Local Similarity: 54.7% Mismatches: 262
Query Match: 53.8% Indels: 57
DB: 2 Gaps: 14
US-09-555-342b-2 (1-1045) x CQ725044 (1-3997)
Qy 1 MetGlyGlnIleGluGlnArg-----ProThrProGlySerArgLeuGlyAla 16
Db 118 ATGGGGAGATAGAAAGAACTACAGAGTCTGACAGCTGAGGAGATCGCTTGGTGGCC 177
Qy 17 ProGluAsnSerGlyIleSerThrLeuGluArgGlyGlnLysProProProThrProSer 36
Db 178 CAGACCCCTGTGGAGTTAGACCTTGGACCTGGGAGACTCTCTTCCCAAGATGCA 237
Qy 37 GlyLeuLeuValSerIleLysIleGlnMetLeuAspAspThrGlnAlaPheGluVal 56

Db 238 GAGAAAGACCTGCACCTCAGAGTAAAGCTGTGACAAACACATGAAATATTGACATT 297
Qy 57 ProGlnArgAlaProGlyLysValLeuLeuAspAlaValCysAsnHisLeuAsnLeuVal 76
Db 298 GAGCTAAATGCGATGCGACAGTATTACTGACCAAGTGTGAAGCGTTTAAACCTGTGA 357
Qy 77 GluGlyAspTyrTrpPheGlyLeuGluPheProAspHisLysValLysIleThrValITrpLeuAsp 96
Db 358 GAATGTACTACTTGGGAAAGAGTTTCAAAATCTAGTCTTACAGTGAATTTGGCTTGAA 417
Qy 97 LeuLeuLysProIleValLysGlnIleArgArgProLysHisValValLysLysPheVal 116
Db 418 CCGATGAACCCATCATTTAGCGAAATACGAAGGCCAAAGATGGTGGCTTGGCTGACT 477
Qy 117 ValLysPhePheProProAspHisLthrGlnLeuGlnGluLeuThrArgTyrLeuPhe 136
Db 478 GTAAATATTTTCCACCTGATCTCGTCACTCAAGAAAGATTTACAGATACTTGT 537
Qy 137 AlaLeuGlnValLysGlnAspLeuAlaGlnGlyArgLeuThrCysAsnAspThrSerAla 156
Db 538 GCCTTGCACTTAAAGAGACCTGTGAAAGAGCGTTTGACCTGTGTCACACACAGCG 597
Qy 157 AlaLeuLeuIleSerHisIleValGlnSerGlnIleGlyAspPheAspGluAlaLeuAsp 176
Db 598 GCCCTTCTACGTCCTCATCTCTGCACTCGGAATAGAGATTCATGAAACGCTGAC 657
Qy 177 ArgGlnHisLeuAlaLysAsnLysTyrIleProGlnGlnAspAlaLeuGluAspLysIle 196
Db 658 CGAGAGCACTTAAAGAGAACGAGTATTTCTCTGGCAGACGACTCCTTGGAAAGATA 717
Qy 197 ValGlnPheHisHisAsnHisIleGlyGlnThrProIleGluSerAspPheGlnLeuLeu 216
Db 718 CTAGAAATTCATAGAAAGACGTTGGGCAACACCTGTGATGCGATTTCCAGTGTCTC 777
Qy 217 GlnIleAlaArgArgLeuGluMetTyrGlyIleArgLeuHisIspProAlaLysAspArgGlu 236
Db 778 GAATTCCTCGAAAGATTGGAATGTACGCACTCAGATTCATGCTTGCACAGGAA 837
Qy 237 GlyThrLysIleAsnLeuAlaValAlaAsnThrGlyIleLeuValPheGlnGlyPheThr 256
Db 838 GGAACCAAGATTCAACGCGAGCTTCCCATGAGTGTACTCGTGTCCAGGGACACACC 897
Qy 257 LysIleAsnAlaPheAsnThrAlaLysValAlaGlyLeuSerPheLysArgLysArgPhe 276
Db 898 AAATCAACACTTTCACCTGCTCCAGAGTCCGTAACTAAGCTTCAAGAGAAAGATT 957
Qy 277 LeuIleLysLeuArgProAspAlaAsnSerAlaTyrGlnAspThrLeuGluPheLeuMet 296
Db 958 CTTATCAAACTTCATCCAGAGCTTCATGACCTTACAGAGACATTTGAATTTTGTG 1017
Qy 297 AlaSerArgAspPheCysLysSerPheTyrLysIleCysValGlnHisAlaPhePhe 316
Db 1018 GTATGTAGATGAATGATTAAGAACTTCTGAAAGATTGTGTGAGATTCACACCTTTT 1077
Qy 317 ArgLeuPheGluGluProLysProLysProLysProValLeuPheSerArgGlySerSer 336
Db 1078 AGACTTTTGACCAACCTAAGCCAAAGCAAAAGCCGCTCTTCAACCGGGCTCTCTCC 1137
Qy 337 PheArgPheSerGlyArgThrGlnLysGlnValLeuAspTyrValLysGlnGlyLys 356
Db 1138 TTCCGATACAGTGGAAAGACTCAGAAACAACTAGTATTTTCCAAAGACAGTGAATG 1197
Qy 357 LysLeuValGlnPheGluArgLysHisSerLysIleHis--SerIleArgSerLeu-- 374
Db 1198 AAGAGAAATTCATATGAAAGAAAGCAACAGCAACCAACGATCCGTTCCAGACTGACT 1257
Qy 375 AlaSerGlnProThrGluLeuAsnSerGlnValLeuGlnGlnSerGlnInsThrSer 394
Db 1258 GCGAGCTTACCAAAACG-----AGC 1278
Qy 395 LeuThrPheGlyGlyGluArgLysArgProGlyGlyGlnSerCysArgArg----- 411

Db 1279 ATCTCATCCCGAGGATTTAGAGACTCTGCTCCCATCTTCAAGGAATGCTTTTAC 1338
 Qy 412 -----GlylySerProlyValSerAlaGlyLysProGly----- 423
 Db 1339 TCGCTCTCTCCCTCCACTGTGTCCCTGTGCGCTGCGAAGTTTAAAGACAGCAGACGC 1398
 Qy 424 -----SerHisProSerProAlaProArgArgSerProAlaGlyLysValSerAlaAsp 441
 Db 1399 TCCCTCACAATATCCCAAGCTTCTCTACGTCAAGAGCTCAGTCCAGAGAGCCAGCGCA 1458
 Qy 442 GlyAlaAlaSerAlaProThrGluGluGluGluValValLysAspArgThrGln 461
 Db 1459 GCATGGCTTGAGAGCCCGAC-----ACACATCG 1488
 Qy 462 SerLysProGlnPro-----GlnProSerThrGlySerLeuThrGlySerPro 478
 Db 1489 GCCACGCTCGAGGCCCCCGACCTCCAGCTGTGTCAGAGCTTTCCAGAAAGAGCTT 1548
 Qy 479 HisLeuSerGluLeuSerValaSerGlnGlyValAlaProAlaAsnValThrLeu 498
 Db 1549 CAGCTTCTCCCTCCAGCCGGAAGAGCCCTGAGTGAAGCCCTGCA-----TTT 1599
 Qy 499 SerProAsnLeuSerProAspThrLysGlnAlaSerProLeuLeuSerProLeuAsn 518
 Db 1600 CAGGTCCCTTGGGCGCCAGCTGAACAGGCTCATCCCACTCTGAGCCCTGTCTCAGT 1659
 Qy 519 AspGlnAlaCysProArgThrArgAspGluAspGluGlyArgArgLysAspArgProThr 538
 Db 1660 GATGCTGGCGGAGCCGGATGATGCTGCGAG--GAGCCAGACACAAAGCCGTGCTGCA 1716
 Qy 539 AspLysAlaTyrPheIleAlaLysGluValSerThrThrGluArgThrTyrLeuLysAsp 558
 Db 1717 GAGAGGCTCTACTTCAATGACAAAGAGATCTGCGTACAGAAAGAACTATCTCAAGAT 1776
 Qy 559 LeuGluValIleThrSerThrPheGlnSerThrValSerLysGluAspAlaMetProGlu 578
 Db 1777 TTGAAGATTATTAACCGTGTGTTCCGCGACGCGAGTGTGAAGAGCGCATGCTGCG 1836
 Qy 579 AlaLeuLysSerLeuIlePheProAsnPheGluProLeuHisLysPheHisThrAsnPro 598
 Db 1837 ACTGTATGACGGTGTCTTCTCCAACTGATCCCATATGAGTTTCCACAGAGCTTTC 1896
 Qy 599 LeuLysGluIleGluGlnArgLeuAlaLeuTrpGluGlyArgSerAlaGlnIleArg 618
 Db 1897 CTCGGGAGGTGAGCAGAGGCTGCGCACTGCGGAAGGCGCTCCAAAGCCACACAAA 1956
 Qy 619 ---AspTyrGlnArgIleGlyAspValMetLeuLysAsnIleGlnLysMetLysHisLeu 637
 Db 1957 GGCAGTCACTCAAGAAATCGGGGACATCTGCTCAGGAACATGCGCGCATTAAGAGGATT 2016
 Qy 638 AlaAlaHisLeuTrpLysHisSerGluAlaLeuGluAlaLeuGluLysAsnGlyIleLysSer 657
 Db 2017 ACCAGCTACTTCCAAATACATGACGAGGCTTAAACGAACTGGAAAGGCTTACCAACGC 2076
 Qy 658 SerArgArgLeuGluAsnPheCysArgAspPheGluLysGluValCysTyrLeuPro 677
 Db 2077 TGTAAAGATTGAGGAGCATCAAGAGATTGAGCTGAGAAAGCTGCTACTTGCCCT 2136
 Qy 678 LeuAsnThrPheLeuLeuArgProLeuHisValGluLeuMetHisTyrLysGlnValLeuGlu 697
 Db 2137 CTCMAACAGTTCTCTGTAAGCCCATCCAGGCTGTGCTGCACTACCGCTGCGTGGCCG 2196
 Qy 698 ArgLeuCysLysHisAspProSerHisAlaAspPheArgAspCysArgAlaAlaLeu 717
 Db 2197 CGCTGTGGGACATTAACGCCCCGGGACCAATGATACCTTACCTGACGCGCTTG 2256
 Qy 718 AlaGluIleThrGluMetValaGlnLeuHisGlyThrMetLysMetGluAsnPro 737
 Db 2257 AAAGCATCAACAAGGTGACCAACACATCAAGACATTTCTATCCGCTGAGAAACCTG 2316
 Qy 738 GlnLysLeuHisGluLeuLysValAspArgLeuIleGlyIleAspAsnLeuValValProGly 757
 Db 2317 CAGAACTTAACGAGCTGCGAGCGGACCTGTGTGGCATAGAAACCTCATTTGCTCTGGC 2376

Qy 758 ArgGluPheIleArgLeuGlySerLeuSerLysLeuSerGlyLysGluLeuGlnIleArg 777
 Db 2377 AGGAGATTATCCGTAGAGGCTGCTTCCACAGCTCAACAAAGAGGCTTGCACGACAGG 2436
 Qy 778 MetPhePheLeuPheAsnAspValLeuLeuTyrThrSerArgGlyLeuThrAlaSerAsn 797
 Db 2437 ATGTTTTTCTGTGTCAGATATGTTGTGTACACAAAGAGAGAGTACAGGACACAGC 2496
 Qy 798 GlnPheLysValHisGlyLeuLeuProLeuTyrGlyMetThrIleGluGluSerGluAsp 817
 Db 2497 CACTTCCGATCCGGGCTCTCTCCCTCCAAAGCATCTGTGTGAGAAAGATGTATAC 2556
 Qy 818 GluTrpGlyValProHisCysLeuThrLeuAlaArgGlyGlnArgGlnSerIleIleValAla 837
 Db 2557 GAGTGTCTGTTCACACACTGTTTCACTACATCTACGCGGCTCAGAAAACATCTGTGGCA 2616
 Qy 838 AlaSerSerArgSerGluMetGluLysTyrValGluAspIleGlnMetAlaIleAspLeu 857
 Db 2617 GCAGCACTCGGCTGAGAAAGAGAGTGAATGATGTGATCTGATCTGAACTCCGATCCAAGCA 2676
 Qy 858 AlaGluLysSerSerSerProAlaProGluPhe-----LeuAlaSerPro 873
 Db 2677 GCCAAGATGGCGGTGACACGCGCTGCATGCCAGCGCGCATCTGTGCACTCGTCC 2736
 Qy 874 ProAspAsnLysSerProAspGluAlaThrAlaAlaAspGlnLysSerGluAspAspLeu 893
 Db 2737 CCC-----AGATCCCCCAAG--GTAATCTTGAGACAGAGATCAGAAAGTATGCT 2787
 Qy 894 SerAlaSerArgThrSerLeuGluArgGlnAlaProHisArgGlyValThrMetValHis 913
 Db 2788 CGGGGTGTCGGAGCTCCCTGAGGGGAGATGGCAGACACGGGCGCAACACATATGCAC 2847
 Qy 914 ValCysTrpHisArgAsnThrSerValSerMetValAspPheSerIleAlaValGluAsn 933
 Db 2848 GTGTGTGTATCCGGAACACAGCGTGCAGGGCAGACACAGTGCACCTGTGAGAAC 2907
 Qy 934 GlnLeuSerGlyAsnLeuLeuArgLysPheLysAsnSerAsnGlyTyrGlnLysLeuTyr 953
 Db 2908 CAGCTTCAAGATATCTGTTAAGAAAGTTCAAAACAGTCAAGCTGCGAGAAAGCTTGG 2967
 Qy 954 ValValPheThrAsnPheCysLeuPhePheTyrLysSerHisGlnAspAsnHisProLeu 973
 Db 2968 GTGCTTTTACCAACTTGTGTTGTTCTTCAAAACATCATCAGGATGATCAACCATG 3027
 Qy 974 AlaSerLeuProLeuGluGlyTyrSerLeuThrIleProSerGluSerGluAsnIleGln 993
 Db 3028 GCCAGCTTCCGCTGTGGCTTACAGCGTACATCCCAAGGAGGCGCATGCGCATACAC 3087
 Qy 994 LysAspTyrValPheLysLeuHisPheLysSerHisValTyrTyrPheArgAlaGluSer 1013
 Db 3088 AAAGACTATGTTTCAAGCTCCAGTTCAAAATCCACGTACTTCTCCGGGTGAGAC 3147
 Qy 1014 GluTyrThrPheGluArgTyrPheGluValIleArgSerAlaThrSerSerAlaSerArg 1033
 Db 3148 AAGTACACATTTGAAAGGTGATGAGGTGATTCAGAGGGCGCAGCTCAGCGGAGAG 3207
 RESULT 12
 BC009153 3719 bp mRNA linear ROD 29-JUN-2004
 LOCUS Mus musculus FERM, RhoGEF and plectristin domain protein 2, mRNA
 DEFINITION (CDNA clone MGC:6304 IMAGE:2655209), complete cds.
 ACCESSION BC009153
 VERSION BC009153.1 GI:14318718
 KEYWORDS MGC.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muridea; Muridae; Murinae; Mus.
 1 (bases 1 to 3719)
 Struhsberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
 Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,

Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marinova, K., Farmer, A.A., Rubin, G.M., Hong, L., Saplinton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Schaefer, T.E., Brownstein, M.J., Uedl, T.B., Tothiyuki, S., Carninci, P., Pringle, C., Raha, S.S., Loguercio, N.A., Peters, G.J., Abramson, R.D., Mulle, J.S., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Feingold, J., Hulton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butler, M.D., Y.S., Krzywicki, M.I., Skalska, U., Smalins, D.E., Schmechel, A., Schein, J.E., Jones, S.J., and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

12477932

2 (bases 1 to 3719)

Strausberg, R.

Direct Submission

Submitted (05-JUN-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgabs-remail.nih.gov

Tissue Procurement: Gilbert Smith, Ph.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LML)

DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center

Center code: BCM-HGSC

Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>

Contact: amg@bcm.tmc.edu

Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Louised, H., Kowls, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati, A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LML at: <http://image.llnl.gov>

Series: IRAC Place: 5 Row: m Column: 20

This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 21704021.

Location/Qualifiers

1. 3719

/organism="Mus musculus"

/mol_type="mRNA"

/strain="FVB/N-3"

/db_xref="taxon:10090"

/clone="MGC:6304.10AGB:2655209"

/tissue_type="Mammary tumor, MMTV-LTR/INT3 model, 5 month old mouse, Taken by biopsy"

/clone_id="NCI_CGAP_Mam2"

/lab_name="DHI08"

/note="Vector: pCMV-SPORT6"

1. 3719

/gene="Farp2"

/note="Synonym: P1r"

/db_xref="GeneID:227377"

/db_xref="MGI:2385126"

99. 3296

/gene="Farp2"

/codon_start=1

/product="FBRM, Rhogef and plectectrin domain protein 2"

/protein_id="AAH09153.1"

/db_xref="GI:14318719"

/db_xref="GeneID:227377"

/db_xref="MGI:2385126"

/translation="MGBIEGTVALPTSGTRLGGQTAIGVSTLEPEQSLSPRMEKIM

RIRVKLDSFVLEFIDIEPKDGOVLLTQWKKHLNLECDYFGLEPFKNVOSYIMLEPM
KPIIROVKRKNVNLRLAAYVFPDPDQLOEBTTRILPALQRLDLEBETCTANTA
ALHISRLQSEIDSDYEDTDREHUKANEYLPNDEKSLKLDTHFGTGPBBDQ
VLIARLLEWYGIREFNADREKTLNLAASHNGVLPQDTKINFTNHSKRVKLSK
RKFFLKLHPEVGPYODTLEPLFGSDENKWKLEAVYHTEFFRPLSDQPKAKV
FSGSSFRYGRGOKOLVDYKDGKMRLEYERHSHSTRTSLALTYLDSKQVSPF
GLTVAISLSANVSPYPPSSSLPCLPKXKQSSSLVPOAPVVKSTAAESSGS
SNGSPRSNHLPGPPVLRGPPSPMSPPSSSKSHSLCPEALDSTAEQAS
PVSPPVSGAGTAMDOEORQKHMPEDATYIAKILATERTYLDLEVTWERS
VILKEEMPAALALFNSIDPYEHRFGLHEVEORLALMEPSSNHLGDIHQRTD
ILLRNKKOLKEFTSYFQDHEVLTLEKATKCKLEAVYHTEFFRPLSDQPKAKV
KPVORLVNHYLLSLRCAHSPGRVADCHALKITEVTELOSLTYLLEMKQT
ELORDVGNVNLAPGREPREGCKHKLTKGLOQMPFLPSMLLYTSASVSGSP
RIGCPVLRGMLVSESENVSVACITVAQCTIVAASTRLSEKMMODLAAVIA
AKTIGSPVPLVGGPPYTRTPRSSDVSLESDDEKGNKSLGNSGRHANTHYVC
YKNTGSVSRADHSAVNSQSLGYLLRKFKNNGKQKLMVETNCLFPYKHODYPLA
SLPLGQSVSLPREADSINHDYVFKLQFSKHVFPFRASVYTEREMWMDVIKASSSPG
RPSFPTQDCSHSPGLEAIRERKACPSPCLDKNL"

ORIGIN

Alignment Scores:

Pred. No.: 3,55e-230 Length: 3719

Score: 2924.00 Matches: 563

Percent Similarity: 70.2% Conservative: 179

Best local Similarity: 53.3% Mismatches: 267

Query Match: 53.5% Indels: 48

DB: 6 Gaps: 9

US-09-555-342b-2 (1-1045) x BC009153 (1-3719)

QY 1 Merglvgluilegu-----GlnArgProthrProglyserArgLeuGlyAla 16

Db 99 ATGGAGAGATGAGAGAGAAATACAGAGCTTTACCACTTCAGAGACCTTGAGGCGC 158

QY 17 ProgluamserglylieserthleugluargglylnlysbProprothrProser 36

Db 159 CAGACCGCCATTGGAGTGAATACCTGGACCAAGAGAGCGCTGCACCAGAGATGACG 218

QY 37 GlylyleuValserilelyslleglMetleuAspAspThrGlnGluAlaPheGluVal 56

Db 219 GAGAAAGCAGATGCGCATCAGAGTGAAGCTGCGACGACCTGATGATTGACATC 278

QY 57 ProgluarglAProglylybValleuLeuAspAlaValGlyAsnHisIleuAsnIleuVal 76

Db 279 GAGCCGAGGTGACGCGGAGGCTTTACTACTCAAGTGGAGAAAGCTTGAATCTCAT 338

QY 77 GluGlyAspTyrrPheglyleuGluPheProAspHisIlelyslIleThrValTyrPleuAsp 96

Db 339 GAATGTGACATCTTGCTGGATTCGAGTTCAAGATGTCAGATCTCACTGATTTGGCTGAA 398

QY 97 LeuLeuLybProIleVallybGlnIleArgArgProLybHisValValValLybPheVal 116

Db 399 CCAATGAAACCAATCACTTAGGCAAGTACGAAACCAAAATAATCGGTGCTCGCTGGCA 458

QY 117 VallybPhePheProProAspHisIleThrGlnGlnGlnGluLeuThrArgTyrrLeuPhe 136

Db 459 GAAATATTTTCCCGCTGACCTCGGTCACTCTCAAGAAAGTACACAGAGTACCTGTTT 518

QY 137 AlaLeuGlnVallybGlnAspLeuAlaGlnGlyArgLeuThrCysAsnAspThrSerAla 156

Db 519 GCCTTGCACTCAAGAGAGACCTCCGAGAGAAAGCGCTGACCTGCACGCGCAACATGCA 578

QY 157 AlaLeuLeuIleSerHisIleValGlnSerGluIleGlyAspPheAspGluAlaLeuAsp 176

Db 579 GCCCTTCTCATATCCACCTTCGACGTGCGAATGGAATATATATATGAAACCTCGAT 638

QY 177 ArgGlnHisLeuAlaLybAsnTyrrIleProGlnGlnAspAlaLeuGlnAspLybIle 196

Db 639 CGAGAACCTTCAAGAGCAATGAAATCTTCCACAGAGAAAGAAATCTAGAAAAGATA 658

QY 197 ValGlnPheHisIleAsnHisIleGlyGlnThrProAlaGluSerAspPheGlnLeu 216

Db 699 CTAGACTTCATCAGAGGACACGCGGCAAGACTCTCGACAGATGACATTTCCAGTGCTT 758


```

Db      2844 CGTAATCAAGTGTGTCTCAGAGCAGACAGTCAGTGTGAGAACCGCTTTCAGGA 2903
Qy      938 AenleuLeuAryLyPheLySerHlaValTyTrpGlnLyLeuTyPValPheThr 957
Db      2904 TATCTGTGAGAAAGTTCAGAACAGTAATGCTGCGCAGAAAGCTCTGGTGTCTTACC 2963
Qy      958 AenPheCysLeuPhePheTyTrpLySerHlaValTyTrpGlnLyLeuTyPValPheThr 977
Db      2964 AACTCTGCTGTGTCTTCTTAATAACATCAGATGATACCCCTGGCCAGGCTTCCA 3023
Qy      978 LeuLeuGlyTyTrpSerLeuThrIleProSerGlnSerGlnLeuGlnIleGlnTyAspTyVal 997
Db      3024 CTATGGGCTTACAGCGCGAGCGCTCCACAGGAGGCTGACAGCATCCCAAAAGCTATGTC 3083
Qy      998 PheLyLeuHlaPheLySerHlaValTyTrpPheArgAlaGlnSerGlnTyTrpPhe 1017
Db      3084 TTCAAGCTCCAAATTCAAATCTCAAGTCTACTTCTCCGGGCTGAGCAAGTACACATTT 3143
Qy      1018 GlnArgTrpMetGlnValIleArgSerAlaThrSerSerAlaSerArgPro 1034
Db      3144 GAAAGGTGATGAGCGTCAATCAAAAGGCGCAGCAGCTCACCGAGGAGACCC 3194

RESULT 13
LOCUS   BC004009
DEFINITION 2632 bp mRNA linear ROD 21-OCT-2003
           Mus musculus FERM, RhoGEF (Arhgef) and pleckstrin domain protein 1
           (Chondrocyte-derived), mRNA (cDNA clone IMAGE:3493093), partial
           cDNA.
ACCESSION BC004009
VERSION   BC004009.1 GI:13278367
KEYWORDS  Mus musculus (house mouse)
SOURCE    Mus musculus
ORGANISM  Mus musculus (house mouse)

REFERENCE
AUTHORS  Straubeberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
          Klausner, R.D., Collins, F.S., Wagner, L., Shennen, C.M., Schuler, G.D.,
          Altschul, S.F., Zeeberg, B., Bueltow, K.H., Schaefer, C.F., Bhat, N.K.,
          Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heide, F.,
          Datchenko, L., Marusik, K., Farmer, A.R., Rubin, G.M., Hong, L.,
          Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
          Scheetz, T.E., Brownstein, M.J., Uebli, T.B., Toshlyuki, S.,
          Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
          Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J.,
          McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
          Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
          Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
          Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,
          Sanchez, A., Whiting, M., Madan, A., Young, A.C., Rodriguez, Y.,
          Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
          Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
          Butterfield, Y.S., Krzywinski, M.I., Skalske, U., Smalins, D.E.,
          Schnerker, A., Schein, J.E., Jones, S.J. and Marra, M.A.
          Generation and initial analysis of more than 15,000 full-length
          human and mouse cDNA sequences
          Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
          12477932
          2 (bases 1 to 2632)
          Straubeberg, R.
          Direct Submission
          Submitted (28-FEB-2001) National Institutes of Health, Mammalian
          Gene Collection (MGC), Cancer Genomics Office, National Cancer
          Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
          USA
          NIH-MGC Project URL: http://mgc.nci.nih.gov
          Contact: MGC help desk
          Email: egabrer-remail.nih.gov
          Tissue Procurement: Gilbert Smith, Ph.D.
          cDNA Library Preparation: Life Technologies, Inc.
          cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

```

DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305
 Web site: <http://www-ehgc.stanford.edu>
 Contact: (Dickson, Mark) mcd@paxil.stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAP Plate: 8 Row: p Column: 11
 This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis.

FEATURES

source

gene

CDs

misc_feature

misc_feature

misc_feature

Location/Qualifiers

1..2632

/organism="Mus musculus"

/mol_type="mRNA"

/strain="Czech II"

/db_xref="taxon:10090"

/clone="IMAGE:3493093"

/tissue_type="Mammary tumor metastatized to lung. Tumor arose spontaneously from a senescent normal mammary (clonal) outgrowth infected with the virus MMTV."

/clone_lib="NCI CGAP_Lu29"

/lab_host="DH10B"

/note="Vector: pCMV-Sport6"

<1..2632

/gene="Farpl"

/note="synonym: Cdep"

/db_xref="GeneID:223254"

/db_xref="MGI:2446173"

<1..1263

/gene="Farpl"

/codon_start=1

/product="Farpl protein"

/protein_id="AAH04009.1"

/db_xref="GI:13278368"

/db_xref="GeneID:223254"

/db_xref="MGI:2446173"

/translations="VNLKNIQCKMLAHLMKHSLEALLETISIKSRRLNHCGRPE
 LQGVLCPLNTFLPLRLHLMLHYKVLRLCKHPPNHADFDRCRALALETFMVLQ
 HGVIMKMFQKLEHLKDLIGDNLVIGRETRILGSLSKGKIQGRMFLFNDY
 LVTLSRGLFASNGFVGOLPLTKYIESEEWGYPHCLTIRGQOSIIIVASSSE
 MEKMMEDIDMAIDLAKEGSPPELLASSPPDKSPDPAADQESDIDASRTISE
 RQAPRGNRTMHTCHRTSVSVNDPSIAVENQSNILRKPRKNSGMCKMTWTFNF
 CLFFYSHDSHPLASLPILGYSLTIPSESNTHQYVRLNFKSHVYTFRASEYTF
 ERMEVIRGATSSAPAHILSHKESHL"

55..309

/gene="Farpl"

/note="RhoGEF: Region: Guanine nucleotide exchange factor for Rho/Rac/Cdc42-like GTPases"

/db_xref="CDD:smat00325"

448..699

/gene="Farpl"

/note="PH: Region: Pleckstrin homology domain. Domain commonly found in eukaryotic signalling proteins. The domain family possesses multiple functions including the abilities to bind inositol phosphates, and various proteins. PH domains have been found to possess inserted domains (such as in PLC gamma, syntrrophins) and to be inserted within other domains. Mutations in Brutons tyrosine kinase (Btk) within its PH domain cause X-linked agammaglobulinemia (XLA) in patients. Point mutations cluster into the positively charged end of the molecule around the predicted binding site for phosphatidylinositol lipids"

/db_xref="CDD:smat00233"

928..1218

/gene="Farpl"

/note="PH: Region: Pleckstrin homology domain. Domain commonly found in eukaryotic signalling proteins. The domain family possesses multiple functions including the abilities to bind inositol phosphates, and various

proteins. PH domains have been found to possess inserted domains (such as in PLC gamma, syntrophins) and to be inserted within other domains. Mutations in Brutons tyrosine kinase (Btk) within its PH domain cause X-linked agammaglobulinemia (XLA) in patients. Point mutations cluster into the positively charged end of the molecule around the predicted binding site for phosphatidylinositol lipids"

/db_xref="CD:smart00233"

ORIGIN

Alignment Scores:

Pred. No.:	6,45e-164	Length:	2632
Score:	2117.00	Matches:	400
Percent Similarity:	97.6%	Conservative:	10
Best Local Similarity:	95.2%	Mismatches:	10
Query Match:	38.8%	Gaps:	0
DB:	6		

US-09-555-342b-2 (1-1045) x BC004009 (1-2632)

```

QY 626 ValMetLeuLysAsnIleGlnGlyMetLysHisLeuAlaHisLeuTrpLysHisSer 645
DB 1 GTCATGCTGAAAAACATTCAGGGCATGAAACACCTGGCGGCCCACTGTGGAAGCACAGT 60
QY 646 GlnAlaLeuGlnAlaLeuGlnAlaGlnGlyIleLysSerSerArgArgLeuGlnAlaAspPheCys 665
DB 61 GAAGCCCTGGAGGCGCCCTGGAGACCTCCATCAAGGGCTCGGGGGGCTGGAGACATTTCTGC 120
QY 666 ArgAspPheGlnLeuGlnLysValCysTyrLeuProLeuAsnThrPheLeuLeuArgPro 685
DB 121 CGAGACTTCGAGCTGCAGAAAGGTGTTCCTCCGCTCAACCTTCCTCCGCGCCCG 180
QY 686 LeuHisArgLeuMetHisTyrLysGlnValLeuGlnLysGlnCysValHisAspPro 705
DB 181 CTCACCGGCTCATGCACTAAAGCATGCTCTGAGAGGCTGTGCAAGCACCAACCA 240
QY 706 SerHisAlaAspPheArgAspCysArgAlaAlaLeuAlaGlnIleThrGlnMetValAla 725
DB 241 AACCAACCGGCATCTTCAGGAGCATCGAAGCTCGCCTGGCGGAGATCAAGAGATGTGGCC 300
QY 726 GlnLeuHisGlnTyrMetLysMetGlnAsnPheGlnLysLeuHisGlnLeuLysLys 745
DB 301 CACCTGCACGCGACCATGATCAAGATGGAAGACCTTCAGAAAGCTGCATGAGCTCAAGAAA 360
QY 746 ArgLeuIleGlnLysLeuAsnLeuValValProGlnArgGlnPheIleArgLeuGlnSer 765
DB 361 GATCTGATCGGCAATGACAACTTGTGATCCAGAGAGGAGATTCATCCGCTGGGAGC 420
QY 766 LeuSerLysLeuSerGlnLysGlnLysGlnLysArgMetPhePheLeuPheAsnAspVal 785
DB 421 CTCAGCAAGCTCTCCGGGAGGGGCTTCACAGGAGCATGTCTTCCTGTTCAACGATGTC 480
QY 786 LeuLeuTyrThrSerArgLysLeuThrAlaSerAsnGlnPheLysValHisGlnGlnLeu 805
DB 481 TTCCTGTATACCAAGCCGGGAGCTGACAGCATCTTAATCATTTAAAGTCCACGACAGCTC 540
QY 806 ProLeuTyrGlnMetThrIleGlnLysSerGlnAspGlnTrpGlnValProHisCysLeu 825
DB 541 CCACTCATGACATGACATGAGATGAGAGATGAGAGATGAGAGATGAGAGATGAGAGATGAG 600
QY 826 ThrLeuArgLysGlnArgGlnSerIleIleValAlaHisLeuSerSerArgSerGlnMetGln 845
DB 601 ACCCTGAGGGGCGACGGGCAATCATGTGTGCTCCAGTTCCCGGTCAAGAGATGAA 660
QY 846 LysTrpValGlnAspIleGlnMetAlaIleAspLeuAlaGlnLysSerSerProAla 865
DB 661 AAATGATGATGAGGACATCCAGATGGCATTCAGCTGCAGAGAAAGCAATGACACCCACC 720
QY 866 ProGlnPheLeuLysSerProProAspAsnLysSerProAspGlnAlaThrAla 885
DB 721 CCCAGACTCTTGCCAGACAGCCCTTCGACAAAGTCCCGGATGAAAGCCACAGCGGCT 780

```

```

QY 886 AspGlnGlnSerGlnAspAspLeuSerAlaSerArgThrSerLeuGlnArgGlnAlaPro 905
DB 781 GACCAAGAGTCAAGAGATGACCTGAGCGGCTCTCCGACATCCCTGGAGGCGGAGGCGCCCT 840
QY 906 HisArgGlnAsnThrMetValHisValCysTrpHisArgAsnThrSerValSerMetVal 925
DB 841 CACCGTGGGCAACATGATGTGACGTGTGTCGTCGATCGGAGACCAAGTCTCCACAGTGTG 900
QY 926 AspPheSerIleAlaValGlnAsnGlnLeuSerGlnLysLeuArgPheLysAsn 945
DB 901 GACTTCAGCATTTGCTGTGAGAAACAGCTCTCTGGGAATCTCTGAGAAAGTTCAAGAAC 960
QY 946 SerAsnGlnTyrGlnLysLeuTrpValValPheThrAsnPheCysLeuPhePheTyrLys 965
DB 961 ACCAATGCTGGCAGAAAGCTGTGAGTGTATTCACCAACTTCTGCTGTTCTTCAAAA 1020
QY 966 SerHisGlnAspAsnHisAspProLeuLysSerLeuProLeuGlnLysSerLeuThrIle 985
DB 1021 TCACACAGAGACATACCCCTTCCAGCCTGCTGTGAGGCTACTCCCTCACCATTC 1080
QY 986 ProSerGlnSerGlnAsnIleGlnLysAspTyrValPheLysLeuHisAspPheLysSerHis 1005
DB 1081 CTTCTGAGTCTGAGAAACATCCACAAGACATGTTCTCAACTGCATTTCAAGTCCAC 1140
QY 1006 ValTyrTyrPheArgAlaGlnSerGlnTyrThrPheGlnArgTrpMetGlnValIleArg 1025
DB 1141 GTGTACTACTCAGGAGCTGAAAGCAATPACATTTGAAAGGTGATGAGATGATCCGA 1200
QY 1026 SerAlaThrSerSerLysSerArgProHisValLeuSerHisLysGlnSerLeuValTyr 1045
DB 1201 AGTGCCACAGCATGACCTCTGAGGCGCCACATCTTGATGACAAAGAAATCTACCTGTAC 1260

RESULT 14
BT024186 4048 bp mRNA linear INV 28-JAN-2006
LOCUS Drosophila melanogaster SD10794 full insert cDNA.
DEFINITION BT024186
ACCESSION BT024186
VERSION BT024186.1 GI:85857423
KEYWORDS FLI CDNA.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 4048)
REFERENCE
  Stapleton,M., Carlson,J., Chavez,C., Frise,E., George,R.,
  Pacleib,J., Park,S., Wan,K., Yu,C. and Ceiniker,S.
  Direct Submission
  Submitted (28-JAN-2006) Berkeley Drosophila Genome Project,
  Lawrence Berkeley National Laboratory, One Cyclotron Road,
  Berkeley, CA 94720, USA
  Sequence submitted by:
  Berkeley Drosophila Genome Project
  Lawrence Berkeley National Laboratory
  Berkeley, CA 94720
  This clone was sequenced as part of a high-throughput process to
  sequence clones from the Drosophila Gene Collection. The sequence
  has been subjected to integrity checks for sequence accuracy,
  presence of a polyA tail and contiguity within 100 kb in the
  genome. Thus we believe the sequence to reflect accurately this
  particular cDNA clone. However, there are artifacts associated with
  the generation of cDNA clones that may have not been detected in
  our initial analyses such as internal priming, priming from
  contaminating genomic DNA, retained introns due to reverse
  transcription of unspliced precursor RNAs, and reverse
  transcriptase errors that result in single base changes. For
  further information about this sequence, including its location and
  relationship to other sequences, please visit our Web site
  (http://www.fruitfly.org) or send email to cdna@fruitfly.org.
  Location/Qualifiers
    1..4048
    /organism="Drosophila melanogaster"
    /mol_type="mRNA"

```

FEATURES

source

CDS

```

/db xref="taxon:7227"
2. _3505
/note="Longest ORF"
/codon_start=1
/product="SD10794p"
/protein_id="ABC6248.1"
/db xref="GI:85857424"
/translation="MSLADMGTAASAGGEGGRHYDLAGTGGSGGHPIVGGGLPGGRH
MLBADIYGLBYQEVSTHRTKWMLEKPMKQVLSLIDPLKFCIYFYPDPQQLS
EYRYFLCQIKDLATGSLQCNNDNTALASYYVQSCDPVEDYFDHTYLSYF
VPMQDAMORKIMENHKHVGQSPAEADLMLETARCELYGMMHFAKQVEGPLN
AVAMGIVFQNTIRINTFSWAKRISFKRKRLVLAHEGYVDYQVAFPEGN
ECKPMKCVENHGFPCATVONTPRKRTLSRSGSPRYSGDKOQIIEFVRNBY
RONRPOSFROGPIASPSOSHTYNSISANPLPIPTAMADNCSQSDSTPBL
TKKADLTDRDNPFGHRSQVTAQVEITYQNKVADSPSDEBACSGAMERCH
HSAVAMDKLNSRSLSPQSPQSWSPSHSHHQRPDPARVHGDANTGTGGINA
MSLDRBEITTPPRYDLTIGSDKSSLSSEAGTYVQAEIOHAROBELGVATA
SHQNGNGNGHTLSTOHDIEAEVKRKPESYFLAKELMTEREYKKDLVLAAT
PROVLSGDVEOLOPFLFELDSLQHNLFRLDIEMHVMQEGSGHEARHIGDVMK
HMAALPYDREVYVTHLDILHCKNDMYEGDERPROVYKEPQKCYPIPIGELLKPLN
RLAHYQILIERLCDYGEENIDYADAAVHHLVSTKGRISQLPDSANVELCELR
DINEQOLVQPHRLIRQGCCLKSKRGLOQRMPELFDLLYGSKSLDPSFRILGV
PVSLLTENAEHNTFSIFGQCAITVAGTAETKLALASKAADIKRPRMQIQ
LTLKNGSSSEEGDLFGLSNGNNSLNSSYGGPLTQOQKLOLQOQOONRTPSR
SYTAHYCMHRAGTGVLGDHLIAEHOLGYLLRKFNSSGMOKLWVFTPSFLCYPK
SYODERPLASLPLIGTYVPRPGODAVQKEPVRKLSFKNHVYFFRAESATYRNLEY
LRSITQTDPRKNVSHAVLGN"

```

ORIGIN

Alignment Scores:

Pred. No.:	4,24e-144	Length:	4048
Score:	1879.50	Matches:	437
Percent Similarity:	54.4%	Conservative:	197
Best Local Similarity:	37.5%	Mismatches:	363
Query Match:	34.4%	Indels:	169
DB:	13	Gaps:	28

US-09-555-342B-2 (1-1045) x BT024186 (1-4048)

```

QY 8 ProThrProGlySerArgLeuGlyAlaProGluAAsenSerGlyIleSerThrLeuGluArg 27
Db 101 CCAATGGCGCGCGGTCTGCCGAGCGCATGACGACCTCCACACACTCTCCGGG 160
QY 28 GlyGlnlyProProThrPro-----SerGlyLeuValSerIleValIle 44
Db 161 GTGGATGGACGCGCTCCACCCGACATCGCGCGCAAAAAGCTGACGTTCGCATC 220
QY 45 GlnMetLeuAAsenThrGlnGluAlaPheGluValProGlnArgAlaProGlyLysVal 64
Db 221 CAGATGCTGACACATCGATACCATGTTCCAAATACAGGTAAAGCATCGGCGCGGTG 280
QY 65 LeuLeuAAsenValCysAAsenHisLeuAAsenValGluGlyAAsenIleValIle 84
Db 281 CTGTTGAACAGGATACCGCTGACGTGAATTACTGAGGCGCACTACTTCGCGTGGAG 340
QY 85 PheProAAsen-----HisLysLysIleThrValIlePheAAsenLeuLysProIle 101
Db 341 TACACAGGAAGTCTCCACACATACCAA-----TACGCGCTGAGTCTCGAAGAGCCATG 394
QY 102 ValLysGlnIleArgArgProLys---HisValValValLysPheValValLysPhe 120
Db 395 AACCGGCAAGTGGCGCTCTCTTATGATCGACAGTGGCTGCGCTTCTGCAAGTCTAT 454
QY 121 ProProAAsenIleThrGlnLeuGlnGluLeuThrArgIleLeuPheAlaLeuGlnVal 140
Db 455 AACCGGCAATCGGCAATGGAAAGAGATACCAAGATATTGTTCTGCTTCAATC 514
QY 141 LysGlnAAsenLeuAlaGlnGlyArgLeuThrCysAAsenThrSerAlaAlaLeuLeuIle 160
Db 515 AAACGAAACCTGGCGCAAGGAGATGTCAGTGAACCAATACACGCGGCGCTTAATATGCA 574
QY 161 SerHisIleValGlnSerGluIleGlyAAsenPhe-----AAsenIleAlaLeuAAsenGlu 178

```

```

Db 575 AGCTACATTGTACAGGGGTATGCGGAGCATCTTCTCCGAGAGATATCCGACCATACT 634
QY 179 HisLeuAlaLysAAsenIleYrIleProGlnGlnAAsenAla---LeuGluAAsenIleVal 197
Db 635 TATCTGCTCCGTATCCCTTGTGTCACCAAGAGACCCACATGACGCAAAAGATATATG 694
QY 198 GlnPheHisIleAAsenHisIleGlyGlnThrProAlaGluAAsenPheGlnIleLeuGlu 217
Db 695 GAGAACCAAAAGACATGTTGGCAATCCCGGAGGCGGACCTTAACCTTTGGAG 754
QY 218 IleAlaArgArgLeuGlnMetYrGlyIleArgLeuHisIleProAlaLysAAsenGluGly 237
Db 755 ACGGCTCGAGGCTGTAGACGTATGAGATGAATAAATGATCCGGGAAGATGCGAAGG 814
QY 238 ThrLysIleAAsenLeuAlaValAlaAAsenThrGlyIleLeuValPheGlnIlePheThrLys 257
Db 815 GTGCGGCTTAATCTGCTGTGCGCCACATGCGGATCACAGTCTTCAGAAATACACCGG 874
QY 258 IleAAsenAlaPheAAsenThrAlaLysValArgLysLeuSerPheLysArgLysPheLeu 277
Db 875 ATCAACACCTTCTGTGGGCTTAAGATACGAAAGATTTCTTCAAGCGCAAGCATCTCG 934
QY 278 IleLysLeuAAsenProAAsenAlaAAsenSerAlaTyrglnAAsenThrLeuGlnPheLeuMetAla 297
Db 935 GTCAATCGATCCAGAGGATATGATATTAACAAGATACCGTGAAGTCTTCTTGGAG 994
QY 298 SerArgAAsenPheCysLysSerPheThrLysIleCysValGlnHisIleAAsenPheArg 317
Db 995 GGTGCGCAACGATGCAAAACCTTCTGGAATAATGCTCGAATAATACGCAATCTTCCGA 1054
QY 318 LeuPheGlu---GluProLysProLysProLysProValLeuPheSerArgLysSer 336
Db 1055 TGCATGCGCTTACAAATATACGCCACGCGCAAACTCGCTCTCTCGGAGGATGTCA 1114
QY 337 PheArgPheSerGlyArgThrGlnLysGlnValLeuAAsenIleValLysGluGlyLysHis 356
Db 1115 TTCGCGTATAGCGGAATAACCGAAGACATATATGCAAGTTCGTCGGAATAATATGTCG 1174
QY 357 LysLysValGlnPheGluArgLysHisSerLysIleHisSerIleArgSerLeuAAsen 376
Db 1175 AAGCGCAAAACCTTCCAAAGTCT-----CAGTATTCGCGCA 1213
QY 377 GlnProThrGluLeuAAsenSerGluValLeuGlnGlnSerGlnGlnIleThrSerLeuThr 396
Db 1214 GGGCAGCTGAATCTGTATAC-----GAAATCAATGCAATACGATATGTA 1260
QY 397 PheGlyGluGlyAlaGluSerProGly-----GlyGln 407
Db 1261 TTCAGATATTACAGCAATCCCTTCTCCCAATTGACATCGCGCATGGGACATATGCA 1320
QY 408 SerCysArgArgGlyLysGluProLysValSerAlaGlyGluProGlySerHisIleProSer 427
Db 1321 TCAATGACAGCATCTGATGACCC-----TTCATGCAAGAAAGGACGCGATACCTT 1374
QY 428 ProAlaProArgArg---SerProAlaGluAAsenLysGlnAla----- 440
Db 1375 GATGCGCGACGACATCTATCGCCACATGCTTTCAAATGCAACAGCCACAGGT 1434
QY 441 -----AAsenIleAlaIleSerAlaProThrGlnGluGlu----- 452
Db 1435 GGAATCTATACAGAAAGAACTATGACGCGGATCGCCCACTCCACGAGGAAGCGGA 1494
QY 453 -----GluValValLysAAsenArgThrGlu 460
Db 1495 ATCTCGCGCGCGGCGGCGGCAACGCAACACATCGCGAGTATGAGCAAACTGAA 1554
QY 460 ngInserLys-----ProGlnProGlnProSerThrGlySerLeuThrGlySe 477
Db 1555 TTGGAACCGCTCACTTTCGCCCAAGGACCGCATGTGTGACT-----TC 1599
QY 477 rProHisIleuSerGluLeuSerValAAsenSerGlnIleGlyValAlaProAlaAAsenVal-- 496

```

```

Db      1600 GCCC-----AGTCACAGACGAGCCACCAACGCGGTCCTCCGATCCGGCTGCTGTGCA 1653
Qy      497 -----ThrluserProAmluserProaph 506
Db      1654 TCCAGCGCATCACATACAGATGGTATATGGAATAATGAAATATGTCGTTGACAG 1713
Qy      506 rlyb-----GlnlaserProleu1 513
Db      1714 GCGCGGTGAATACCACTCCACCAACGCGGTACAGCTCAGCTTAGATCGGATAGATC 1773
Qy      513 eSerProleuLeuasnApGlnla----- 521
Db      1774 TTGTCATCATATCCCGATCTGAGGACAGGAACATATGACGTATATACAGGACAAATTCAGCA 1833
Qy      521 ----- 521
Db      1834 CGCTAAGACAGAGAGCTAGCCACAGGCGTTCGCCACCGCTCGCACCAAGATGGAAACGG 1893
Qy      522 -----CysProArgThrAspAspGluAspGluArgArgly 534
Db      1894 AAACGGAAATGGGCATATCCCTGACACACACACAGATATTGAGGACGAAAGTGAAGG 1953
Qy      534 sArgPheProThrAspLyalaTyPheilealalyGluValaSerThrThrluArgTh 554
Db      1954 GAAATGGCCACCGAGCCCGCTACTTTTGGCCAAAGAGCTGTGATGACGGAACGTAC 2013
Qy      554 rTyPLeuValAspLeuGluValleThrsertPrpHeGlnSerThValSerlyGluVal 574
Db      2014 GTACAAAGAGATCTGATGTGCTGAACACACACCTTCGACAGGTGCTGAGTCTCGGGA 2073
Qy      574 palmetProGluValaleuYserLeuilePheProAspPheGluProleuHlyAsp 594
Db      2074 TGG-----GAGCAGGTGACGAGCGCGT-----TTCAGAGCTGCTGACCTCCGCGCCAGCA 2124
Qy      594 eHlsthAenPheLeuYsgluileGluValaLeuTrpGluValArgSerAe 614
Db      2125 CCACAACTCTCTCTCTGCGACATTCGAAACCCGATCTGTCAGTGGAGGAAAGAGAGG 2184
Qy      614 nlaGlnleArgAspTyrglnArgleGluValMetLeuLysAsnleGlnGlyMe 634
Db      2185 C-----CACGAAGCCACCGCATCGAGAGCGTATATAGACACATGCGCGCTCT 2235
Qy      634 clybHlaleuValaHlaleuTrpLysHlaserGluValaleuGluValaLeuGluVal 654
Db      2236 GCCCATCTACAGATGAGTACGTCAGACGACCTGACATCTTGCATGTAACACGAT 2295
Qy      654 ylleYsSerSerArgArgleuGluValAspPheCyAspAspPheGluLeuGlnlyValCy 674
Db      2296 GTACGAAGCGCATGAACGCTTCGTCAGGTGTACAGAGAGTTCGAGCAGAAAGGTTTG 2355
Qy      674 sTyPLeuProLeuAsnThrPheLeuValArgProLeuHlAsxLeuMetHlTyPlysgl 694
Db      2356 CTATCTACCATCGCGCACTTCTACTGAAGCCCTTCAACCGCGCTGTTCCTACCACT 2415
Qy      694 nValLeuGluArgLeuCylybHlAspProSerHlAslAspPheArgAspCyAsp 714
Db      2416 GATCTTAGAGCGGCTGCGACACTATAGGAGAGAGACATATGACATGCGCATGTAT 2475
Qy      714 galalaaleuValGluileThrluMetValalaGlnleuHlGlyThrMetlelybMe 734
Db      2476 GGCCTGCGACACACTTCTGCTTCGACGACCAAGGATATTGCTCCAGCTCCGACTC 2535
Qy      734 tGluAnPheGlnlybLeuHlGluLeuLyslybAspLeuileGlyleAspAnleuVal 754
Db      2536 TGCAAACTTGTGAGAGCTGTGTAGACTGCAACGCGAC--ATCAACTTCGACAGCTTGGT 2592
Qy      754 lvalProGlyArgGluPheHlAsxArgleuGlySerLeuSerlybLeuSerGlylysglyle 774
Db      2593 TCAAGCCCATCGCGCTCATCCGCCAGGATGCTTCGAAAGACTCCAAAGCGCGGTCT 2652
Qy      774 uGlnGlnArgMetPhePheLeuPheAnAspValleuLeuTyThSerArgGlyLeuTh 794
Db      2653 GCAGCAGAGATGTCTTCTTGTGTTCCAGCTGCTCTACGCTTCGAAGTCC--CC 2709

```

```

Qy      794 rAlaserAenGlnPheLyValHlGlyGlnLeuProLeuTyrglyMetThrlleGluG 814
Db      2710 GCTGACACAGACTTCGCCATTTTAGCCACAGCTTCTGTGCGTTACTATTCTACCGAGAA 2769
Qy      814 uSerGluAspGluTyPglyValProHlCySerLeuThrlAsxArglyGlnArgGlnSer11 834
Db      2770 CGCCGAGCAC-----AACACTTCTCATCTTCGCGGAGACAGTGTCCAT 2814
Qy      834 eilleValalaaserSerArgSerGluMetGluTySTPrValGluAspIle---GlnMe 853
Db      2815 CACAGTGAAGTGGCGGACCGCTGAAGAAAGCCCTTGCTGGCCGAGCTTCCAAAGG 2874
Qy      853 lAlaIleAspLeuAlaGluLyb----- 860
Db      2875 GGCCTGGGATATTAAAAACCGCCCAATGCAACTCCAGCTAACGACGCTCAAGAA 2934
Qy      861 -----SerSerSerProAlaProGluPheLeuAlaSerSerProAspAnlySg 878
Db      2935 CTGCAAGCTCTCGAGAGAGGCGCTGACTATTGCGTGTGACAGACGAGAAACAGCAG 2994
Qy      878 rProAspGluValaThrlAlaAsp-----GlnGluSerGluAspAspSerAl 895
Db      2995 CTTGAACAGACAGTGAACGCGCGCGCTCCGCTGACTACGACGACGACAGATTGACGCT 3054
Qy      895 aSerArgThrsertLeuGluArgGlnAlaProHlAsxArglyAsnThrMetValHlValCy 915
Db      3055 GCAGCAACAGCAGCAAAACAGACGAGCATCCGACAAATAGCGCTGACAGCTCG 3114
Qy      915 sTyPHisArgAnThrservalSerMetValAspPheSerlleAlaValGluAnGlnle 935
Db      3115 TTGGCATCTGTGCGCGCACCGTGAAGCTGGCGCATCACTGATAGCCGCGAGACACAGT 3174
Qy      935 uSerGluAnleuLeuArgLybPheLybAsnSerAnGlyTyTrpGlnlybLeuTyValVa 955
Db      3175 GTCCGGTATCTGCTGCGAAAGTTCAAGATAGCTCGCTGCGACAAAGCTCTGGGTGT 3234
Qy      955 lPheThraenPheCylybPhePheTyrglySerHlAsxGlnAspAnHlProleuAlaSe 975
Db      3235 GTTCACAGTCTTGTGCTGTACTTTTCAAGAGCTACAGAGAGTTCGACATGCTGAG 3294
Qy      975 rleuProleuLeuGlyTyrglySerleuThrlleProserGluSerGluAnHlGlnlybAs 995
Db      3295 TCTGCGCTTATGGTGTGTACAGGCTTCCTGCTGCCACCAAGAGCCGCTGCAAGAGA 3354
Qy      995 pTyrgValPheLybLeuHlAsxPheLybSerHlValTyTyTrpAsxArgAlGluSerGluTy 1015
Db      3355 GTTCGCTTCAAGCTTCTTTAAGAACAGCTACTTCTTCGCGGAGAAAGTCCGCA 3414
Qy      1015 rThPheGluArgTyPmetGluValleArgSerAlaThrserSerAlaserArgProHl 1035
Db      3415 CACCTAACAAAGATGCTGAGAGTACTGCGACGACCAACCCAAACCCAGACTTCAAGAA 3474
Qy      1035 sValleuSerHis 1039
Db      3475 CGTACACAGCCAT 3487

RESULT 15
BC027077 2431 bp mRNA linear ROD 21-OCT-2003
LOCUS Mus musculus FERM, RHOGEF (Arhgef) and plectrokin domain protein 1
DEFINITION (chondrocyte-derived), mRNA (cDNA clone IMAGE:5010682), partial
cde.
ACCESSION BC027077
VERSION BC027077.1 GI:20071584
KEYWORDS Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 2431)
AUTHORS Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,

```

Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heien F., Diatchenko L., Maruina K., Farmer A.A., Rubin G.M., Hong L., Scapellato M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Uedl T.B., Toohilyki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J., Boeck S.A., McKean P.J., McKernan K.D., Malek J.A., Gunatane P.H., Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fehey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Gilmwood J., Schmutz J., Myers R.M., Butterfield Y.S., Krzywinski M.I., Skalska U., Smalins D.E., Scherneck A., Schein J.E., Jones S.J. and Marra M.A.
Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

JOURNAL
PUBMED
15477932
2 (bases 1 to 2431)
Strausberg R.
Direct Submission
Submitted (04-APR-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK
COMMENT
NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
Contact: MGC help desk
Email: gcgaps-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: amg@bcm.tmc.edu
Gunnarsson, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Louised, H., Kowls, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanaavati, A.N., Gibbs, R.A.

FEATURES
Source
Location/Qualifiers
1..2431
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N-3"
/db_xref="taxon:10090"
/clone="IMAGE:5010682"
/cbase_type="mammary tumor. MMTV-LTR/INT3 model. 5 month old mouse. Taken by biopsy."
/clone_id="NCI CGAP Mam2"
/lab_host="DH10B"
/note="vector: pCMV-SPORT6"
1..2431
/gene="Farp1"
/note="Synonym: Cdep"
/db_xref="GeneID:223254"
/db_xref="MGI:2446173"
1..1065
/gene="Farp1"
/codon_start=1
/product="Farp1 protein"
/protein_id="AAH27077.1"
/db_xref="GI:20071585"
/db_xref="GeneID:223254"

CDS

/db_xref="MGI:2446173"
/translation="YKVLRLCKHPNPNDPRDCAALAEITMVAOLGTMKME
NPKQLHKLKOLIGDNLTPGRRFRLGSLKSGGLQORPFLPNDLTLTSL
TAQNQFVHQLPLVYGTITSESEEWVPECLTLRGROSIIVAASRSMEKMEI
OMADLAEKNSGPTPELASEPDKNSPDATADQSEEDLSASRSLEORPHEN
TQNHVCHRSTSVMSVDFSAVENQSGNLRRFKNSGKGMVFTNCLPEYKSH
QDSHPPLASLILPLGSLTIPESINIKDVFYKLFKSHVYFPAESYTERMWEVIR
SATSSRAHILHSHKSHLY"
250..501
/gene="Farp1"
/note="PH: Region: Pleckstrin homology domain. Domain
commonly found in eukaryotic signalling proteins. The
domain family possesses multiple functions including the
abilities to bind inositol phosphates, and various
proteins. PH domains have been found to possess inserted
domains (such as in PLC gamma, syntrophins) and to be
inserted within other domains. Mutations in Brutons
tyrosine kinase (Btk) within its PH domain cause X-linked
agammaglobulinemia (XLA) in patients. Point mutations
cluster into the positively charged end of the molecule
around the predicted binding site for phosphatidylinositol
lipids"
/db_xref="CD:smart00233"
730..1020
/gene="Farp1"
/note="PH: Region: Pleckstrin homology domain. Domain
commonly found in eukaryotic signalling proteins. The
domain family possesses multiple functions including the
abilities to bind inositol phosphates, and various
proteins. PH domains have been found to possess inserted
domains (such as in PLC gamma, syntrophins) and to be
inserted within other domains. Mutations in Brutons
tyrosine kinase (Btk) within its PH domain cause X-linked
agammaglobulinemia (XLA) in patients. Point mutations
cluster into the positively charged end of the molecule
around the predicted binding site for phosphatidylinositol
lipids"
/db_xref="CD:smart00233"

misc_feature

ORIGIN
Alignment Scores:
Pred. No.: 1,54e-136 Length: 2431
Score: 1784.00 Matches: 338
Percent Similarity: 97.7% Conservative: 8
Best Local Similarity: 95.5% Mismatches: 8
Query Match: 32.7% Indels: 0
DB: Gaps: 0

US-09-555-342B-2 (1-1045) x BC027077 (1-2431)

QY 692 TTYLGGINValleuGluatGLeuCySLyPHiBProProSerhIbAlaapphearg 711
Db 1 TATTAAGCATGTCTCGAGAGCTGTGCACACACCAACCAACGCGGACTTCAGG 60
QY 712 ASPCyAaRgAlaIaleuAlaGluIleThrGluMeRValAlaGluMeuHIsGlyThMet 731
Db 61 GACTGCAAGAGCTGGCGGAGATACAGAGTGTGCACAGCTGCATGGTACCATTG 120
QY 732 IlleSwetGluAspPheGlnTylsLeuHIsGlyLeuLysAlaPheLleGlyIleasp 751
Db 121 ATCAAGATGAGAACTTCAGAAAGCTGCATGACTCAAGAAAGATCTGATCGGATTGAC 180
QY 752 AsnLeuValAlaProGlyATgTgUphelIaargLeuGlyserLeuSerGly 771
Db 181 AATCTGTGACCCAGAGAGGAGTTCATCGGCTGGCGCCTCAGCAAGCTTCGCGG 240
QY 772 LyeGlyLeuGlnIaRgMetPhePheLeuPheAsnAspValLeuLeuTyThrSerArg 791
Db 241 AAGGGGCTTCAAGCAGGCAATGTTCTTCTTCAAGATGTTCTTGCTGTATACAGCGG 300
QY 792 GlyLeuThrIaSerAngInPheLysValHIsGlyGluLeuProLeuTyrglyMetThr 811
Db 301 GGACTGACGGCATCTATCATGTTTAAAGTCCACGAGACGCTCCACATCTATGCGATGACG 360

```
QY 812 ILeGIuSerGIuAspGIuTrpGIuValProHieCySeruThrLeuArgGIuIArg 831
DB 361 ATCGAGAGAGGTAGAGAGGTGGGTGTGCCCCACTGCTGACCTTGCGGGGCACGG 420
QY 832 GInSerIleIleValAlaIaSerSerArgSerGIuMetGIuIuYsTrpValGIuAspIle 851
DB 421 CAATCTATCATCGTGGCTGCCAGTCCCGGTCAAGATGGAATAATGATGAGAGACATC 480
QY 852 GInMetAlaIleAspLeuAlaGIuIuYsSerSerSerProAlaProGIuPheLeuAlaSer 871
DB 481 CAGATGCGCCATTGACCTGCGCAGAGAAAGCAATGACCCGCCGAGCTCTTGCCAGC 540
QY 872 SerProProAspAsnIuYsSerProAspGIuAlaThrAlaAlaAspGIuSerGIuAsp 891
DB 541 AGCCCCCTTACAAACAAAGTCCCGGATGAAGCCACAGCGGTGACAGAGATCAGAGAT 600
QY 892 AspLeuSerAlaSerArgThrSerLeuGIuArgGIuAlaProHieArgGIuAsnThrMet 911
DB 601 GACCTGAGTGCCTCTCGGACATCGCTGAGCGCCAGGCCCTCACCCTGGCAACAGATG 660
QY 912 ValHieValCySerTrpHieArgAsnThrSerValSerMetValAspPheSerIleAlaVal 931
DB 661 GTGCACGTGTGCTGGCATCGACACACCATGTCTCCATGTGTGACTTCAGCATTTGCTGTG 720
QY 932 GIuAsnGIuIuSerGIuAsnLeuLeuArgIuYsPheIuAsnSerAsnGIuTrpGIuIuYs 951
DB 721 GAGAACCAAGCTCTTGAGAACTTGCTGAGAAAGTTCAAGAACAGCAATGGCTGGCAGAG 780
QY 952 LeuTrpValIaPheThrAsnPheCyLeuPhePheTyrIuYsSerHieGIuAspAsnHie 971
DB 781 CTGTGGGTGTATTCACCACTTGCTGTCTTCTTACAAATCACACGACGACAGTCAC 840
QY 972 ProLeuAlaSerLeuProLeuLeuGIuTyrSerLeuThrIleProSerGIuSerGIuAsn 991
DB 841 CCCCTTGCCAGCGCTGCTGTGGGCTACTCCCTCACCATCCCTCTGAGTGTGAGAAC 900
QY 992 ILeGIuIuYsAspTyrValPheIuYsLeuHiePheIuYsSerHieValTyrTyrPheArgAla 1011
DB 901 ATCCACAGAGACTATGTGTTCAGAGCTGCATTTCAGTCCACGCTGACTTCAAGGCT 960
QY 1012 GIuSerGIuTyrThrPheGIuArgTrpMetGIuValIleArgSerAlaThrSerSerAla 1031
DB 961 GAAAGCGAATACACATTGAAAGGTGATGGAAGTATCCGAAGTGCCACCACTCAGCC 1020
QY 1032 SerArgProHieValIuSerHieIuYsGIuSerLeuValTyr 1045
DB 1021 TCTGGGCGCCACATCTTGAGTCAACAAATCTCACCTGTAC 1062
```

Search completed: May 20, 2006, 05:44:36
Job time : 14331 secs

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: May 20, 2006, 01:25:21 / Search time 1473 Seconds
(without alignments)
7419.552 Million cell updates/sec

Title: US-09-555-342B-2
Perfect score: 5463
Sequence: 1 MGIEBRPPTGSRGAPENS.....SATSSASRPVHSHKESLVY 1045

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 5244920 seqs, 3486124231 residues
Total number of hits satisfying chosen parameters: 10489840

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+.p2n.model -DEV=xlh
-Qc/abses/ABSSWEB.spool/US09555342/runcat.19052006.121452.13173/app.query.fasta.1
-DB=N.GeneSeq -QMT=fastcap -SUPFIX=p2n.rng -MINMATCH=0.1 -LOOPEXT=0 -LIST=45
-UNITS=bits -START=1 -END=-1 -MATRIX=blonum62 -TRANS=human40.cdd -LIST=45
-DOCCALIGN=200 -THR.SCORE=pcr -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=prco -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abses05h
-USER=US09555342.OCGN.1.1.761.0runcat.19052006.121452.13173 -NCPU=6 -ICPU=3
-NO.MMP -NEG.SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG -DEV.TIMEOUT=120
-MARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_GeneSeq.8:.*
1: geneseqn1980a:.*
2: geneseqn1990a:.*
3: geneseqn2000a:.*
4: geneseqn2001a:.*
5: geneseqn2001b:.*
6: geneseqn2002a:.*
7: geneseqn2002b:.*
8: geneseqn2003a:.*
9: geneseqn2003b:.*
10: geneseqn2003c:.*
11: geneseqn2003d:.*
12: geneseqn2004a:.*
13: geneseqn2004b:.*
14: geneseqn2005a:.*
15: geneseqn2006a:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5463	100.0	3442	2	AAK79183
2	5463	100.0	3442	15	AEE03370
3	5463	100.0	5028	14	AEB22714

4	5459	99.9	4687	3	AAA08582	aa08582 Human cys
5	5431	99.4	3187	5	AA64760	aa64760 DNA encod
6	2941.5	53.8	3997	12	ADG84807	adg84807 Human tun
7	2941.5	53.8	3997	13	ADR25675	adr25675 Breast ca
8	2941.5	53.8	3997	13	AD644452	ad644452 Human KIA
9	2498	45.7	3094	3	AAC89992	aac89992 Human pan
10	1150.5	21.1	3200	12	ADG63863	adg63863 Novel hum
11	1006.5	18.4	1718	4	AA527016	aa527016 cDNA enco
12	1006.5	18.4	1718	10	ADB93194	adb93194 Human cDN
13	964	17.6	582	10	ACD92278	acd92278 Human col
14	964	17.6	582	10	ACD94832	acd94832 Human col
15	913.5	16.7	1747	4	ABL03867	abl03867 Drosophi1
16	905.5	16.6	1220	4	ABL02335	abl02335 Drosophi1
17	891.5	16.3	4150	4	ABL03865	abl03865 Drosophi1
18	754.5	13.8	2277	10	ADCT1321	adct1321 Human col
19	754.5	13.8	3275	13	ADRO7347	adro7347 Full leng
20	748.5	13.7	3309	13	ADRI4498	adri4498 Human NF-
21	748.5	13.7	3309	14	ADY16637	ady16637 DNA encod
22	748.5	13.7	3309	14	ADY15739	ady15739 DNA encod
23	748.5	13.7	3309	14	ADZ22659	adz22659 Erythrocy
24	744	13.6	4543	10	ABT42346	abt42346 Toxicity
25	743.5	13.6	3387	13	ADRI4396	adri4396 Human NF-
26	739.5	13.5	3770	13	ADRO8260	adro8260 Full leng
27	735.5	13.5	14917	6	AA023338	aa023338 Human lun
28	727	13.3	6305	10	ADC31979	adc31979 Human nov
29	727	13.3	6305	3	AA081980	aa081980 Human nov
30	724	13.3	3620	10	AA08581	aa08581 Human cys
31	724	13.3	3620	10	ADJ56474	adj56474 Murine CD
32	724	13.3	4336	10	ADH28737	adh28737 Human chr
33	724	13.3	4336	14	ADY14625	ady14625 DNA encod
34	724	13.3	4336	14	ADY14627	ady14627 DNA encod
35	724	13.3	4336	14	ABG2823	abg2823 Breast ca
36	717.5	13.1	6274	10	ADC30054	adc30054 Human nov
37	717	13.1	637	14	ACLS5653	ac15653 Human col
38	716.5	13.1	2410	8	ACC44336	acc44336 Gene enco
39	716.5	13.1	4338	14	AEA19657	aea19657 Novel hum
40	715	13.1	671	13	ADQ52709	adq52709 Novel can
41	714	13.1	435	14	ACI61462	aci61462 Human col
42	713	13.1	2656	14	AEA19659	aea19659 Novel hum
43	713	13.1	2867	12	ADP10335	adp10335 Reference
44	713	13.1	2867	13	ADR90365	adr90365 Human ful
45	713	13.1	2867	13	ADP54688	adp54688 Human PRO

ALIGNMENTS

RESULT 1	AAK79183	AAK79183 standard; DNA; 3442 BP.
ID	AAK79183	
XX	AAK79183	
AC	17-AUG-1999 (first entry)	
DT	17-AUG-1999 (first entry)	
XX		
DE	Human chondrocyte-derived gene CDEP.	
XX		
KW	Differentiation; human; foetal; chondrocyte; ezrin-like domain; cancer;	
KW	DBI homology domain; plectroin homology domain; rheumatoid 'arthritis;	
KW	drug; ss.	
XX		
OS	Homo sapiens.	
XX		
XX	WO9928458-A1.	
FN	10-JUN-1999.	
PD		
XX		
PF	27-NOV-1998; 98MO-JP005348.	
XX		
PR	27-NOV-1997; 97JP-00342060.	
XX		
PA	(CHUS) CHUGAI SEIYAKU KK.	
XX		
PI	Kato Y, Kawamoto T, Koyano Y;	

XX WPI: 1999-371117/31.
DR P-PSDB; AAY07482.
XX Protein CDEP expressed in differentiated chondrocytes, and gene encoding
PT it.
XX
PS Claim 5; Fig 1; 59pp; Japanese.
XX
CC This sequence represents the coding region for a protein (CDEP) expressed
CC in differentiated human foetal chondrocytes, which contains an ezrin-like
CC domain, a Dbl homology (DH) domain and a pleckstrin homology (PH) domain.
CC The nucleic acid or protein can be used in the investigation and
CC treatment of cancers and arthritic diseases (including chronic rheumatoid
CC arthritis), or for screening of candidate anticancer drugs
XX
SQ Sequence 3442 BP; 864 A; 952 C; 927 G; 699 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 0 Length: 3442
Score: 5463.00 Matches: 1045
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
Gaps: 0
US-09-555-342B-2 (1-1045) x AAX79183 (1-3442)

QY 1 MercGlyGluIleGluGlnArgProThrProGlySerArgLeuGlyAlaProGluAAsnSer 20
DB 49 ATGGGAGAAATAGACGAGAGCCGACCCGAGATCAGACTGGGGGCCCGGAAATTCG 108
QY 21 GATLISerThrLeuGluArgGlyGlnIleArgProProThrProSerGlyLysLeuVal 40
DB 109 GGGATGATACCTTGGACGTGGACAGAACCGCCCCAACCTTCAGAAATCTGTG 168
QY 41 SerLleLysIleGlnMetLeuAspAspThrGlnIleAlaPheGluValProGlnArgAla 60
DB 169 TCATCAAAATCCAGATGCTGGATGACACCCAGAGGCATTGTAAGTTCCAAAGAGCT 228
QY 61 ProGlyLysValLeuLeuAspAlaValCysAsnHisLeuAsnLeuValGluGlyAspTyr 80
DB 229 CCGGGAGAGGTGCTGCGATGACATTTGGACCACTCAACTCGGGAGAGGTGACTAT 288
QY 81 PheGlyLeuGluPheProAspHisLysLysIleThrValTyrLeuAspLeuLeuLysPro 100
DB 289 TTGGGCTCGAGTTTCTCTGATCAAAAGATCACGGTGGCTGGATCTCTTAAACCC 348
QY 101 IleValLysGlnIleArgArgProLysHisValValLysPheValValLysPhePhe 120
DB 349 ATGTGAAACAGATTAGAGGCCAAGCACTTGTGTTAAGTTTGTGGAAATCTTT 408
QY 121 ProProAspHisThrGlnLeuGlnGluLeuThrArgTyrIleuPheAlaLeuGlnVal 140
DB 409 CCGCTGCACACACAACTCCAAAGAAACTCAAGAGTACTGTTCGGCTGCAGGTG 468
QY 141 LysGlyAspLeuAlaGlnGlyArgGlyLeuThrCysAsnAspThrSerAlaAlaLeuLeuIle 160
DB 469 AAGCAGAGCTTGGCTCAAGGACAGGTGACGTGATACACACAGCGCACTCTTGATT 528
QY 161 SerHisIleValGlnSerGluIleGlyAspPheAspGluAlaLeuAspArgGlnHisLeu 180
DB 529 TCACACATTTGCAATCTGAGATTGGGGATTGATGATGAACTTTGGACAGAGCACTTA 588
QY 181 AlaLysAsnLysTyrIleProGlnGlnAspAlaLeuGluAspLysIleValGluPheHis 200
DB 589 GCAAAATTAATCAATCACTCAGCAAGACGACTAAGACAAATTCGGGAATTCAC 648
QY 201 HisAsnHisIleGlyGlnThrProAlaGluSerAspPheGlnLeuLeuGlnIleAlaArg 220
DB 649 CATTAACCACTTGGACAAACACAGCAAGATCAAGATTTCCAGCTCTTGAAGATTGGCCGT 708
QY 221 ArgLeuGluMetTyrGlyIleArgLeuHisProAlaLysAspArgGluGlyThrLysIle 240

DB 709 CGCGTAGAGATGATGAAATCCGGTGCACCGGCGCAAGACAGGAGCAAGCAAGATC 768
QY 241 AsnLeuAlaValAlaAsnThrGlyIleLeuValPheGlnGlyPheThrLysIleAsnAla 260
DB 769 AATCTGGCCGTGGCCAAACGGGAATCTTAGTGTTCAGGGTTTCATAGATCAATATGCC 828
QY 261 PheAsnTrpAlaLysValArgLysLeuSerPheLysArgLysArgPheLeuIleLysLeu 280
DB 829 TTCACTGGGCCAAGGTGGGAAGCTGAGCTTCAAGAGAAAGGCTTTCATCAAGCTC 888
QY 281 ArgProAspAlaAsnSerAlaTyrGlnAspThrLeuGluPheLeuMetLaseArgAsp 300
DB 889 CGGCCAAGTCCAAATGTGCTGATCAGATACGATTCCTGGAATTCCTGATGGCAGTCCGGAT 948
QY 301 PheCysLysSerPheThrLysIleCysValGluHisAlaPhePheArgLeuPheGlu 320
DB 949 TTCTGCAGTCTCTTCGAAAAATCTGTGTGAACATCAATGCTTCTTAGACTTTTGA 1008
QY 321 GluProLysProLysProLysProValLeuPheSerArgGlySerSerPheArgPheSer 340
DB 1009 GAGCCCAAAACAAAGCCCAAGCCCGTCTCTTAGCCGGGGGTCAATCATTTCCGTTCAAG 1068
QY 341 GlyArgThrGlnLysGlnValLeuAspTyrValLysGluGlyGlyHisLysLysValGln 360
DB 1069 GGTCCGAGCTCAGAGACAGGTTCTCGACTATGTTAAAGAGAGGACATTAAGAGGTGCAG 1128
QY 361 PheGluArgLysHisSerLysIleHisSerIleArgSerLeuAlaSerGlnProThrGlu 380
DB 1129 TTTGAAAGAGACACGAAGATTCATTCATCCGAGGCTTGCTTCACAGCTTACAGAA 1188
QY 381 LeuAsnSerGluValLeuGluGlnSerGlnGlnSerThrSerLeuThrPheGlyGluGly 400
DB 1189 CTGAATTCGGAAGTGTGAGAGAGTCTCAGCAGAGACCAAGCTTATTTGGAGAAAGGT 1248
QY 401 AlaGluSerProGlyGlyGlnSerCysArgArgGlyLysGluProLysValSerAlaGly 420
DB 1249 GCCGAATCTCCAGGGGGCCAGAGCTCCGCGGAGAAAGAACCGAAGGTTTCCGGCGG 1308
QY 421 GluProGlySerHisProSerProAlaProArgArgSerProAlaGlyAsnLysGlnAla 440
DB 1309 GAGCCGGGGTCCGACCCGAGCCCTGCGGAGAGAGAACCCCGGGGTAAACAGCAGCG 1368
QY 441 AspGlyAlaAlaSerLysIleProThrGlnGluGlnGluValAlaLysAspArgThrGln 460
DB 1369 GACGAGACCGGCTCGCGCCACGAGAGAAAGAGAGAGTGTTAAGATAGACCCAG 1428
QY 461 GlnSerLysProGlnProProGlnProSerThrGlySerLeuThrGlySerProHisLeu 480
DB 1429 CAGAGTAAACCTCAGCCCGGACGCMACACAGAGCTCCCTGACTGCGAGTCTCACTT 1488
QY 481 SerGluLeuSerValAsnSerGlnGlyValAlaProAlaAsnValThrLeuSerPro 500
DB 1489 TCCGAGCTGTGTGAATCTGCAAGGGGAGAGGCCCTTGCCAACTGATCTGTCTCC 1548
QY 501 AsnLeuSerProAspThrLysGlnAlaSerProLeuIleSerProLeuLeuAsnArgGln 520
DB 1549 AACCTAGGCCGACACCAAGCAGGCTCTCTCTTATCAGCCGCTGCTGAATGACAG 1608
QY 521 AlaCysProArgThrAspAspGluAspGluGlyArgLysArgPheProThrAspLys 540
DB 1609 GCCTGCCCCCGGACGACATGAGATGAGGGCCGAGAGAAAGATTTCCCACTGATTAA 1668
QY 541 AlaTyrPheIleAlaLysGluValSerThrThrGlnArgThrTyrLeuLysAspLeuGlu 560
DB 1669 GCGTACTTAATAGCTAAGAAAGTGTTCACCAAGCCGAACATATTTGAAGATTCGAA 1728
QY 561 ValIleThrSerTyrPheGlnSerThrValSerLysGluAspAlaMetProGluAlaLeu 580
DB 1729 GTTATCATCTTCGAGTTTCAAGACACAGTGAAGAAAGAGAGCCATGCGGAAGACAGT 1788
QY 581 LysSerLeuIlePheProAsnPheGluProLeuHisLysPheHisThrAsnPheLeuVal 600

Db 1789 AAAAGTCATATATCCGAAATTTTGAACCTTTGGCAAAATTTTCACTAATTTTCTCAG 1848
 Qy 601 G|U|I|E|G|U|N|A|G|U|A|L|E|U|T|P|G|U|G|Y|A|G|S|E|R|A|N|A|A|G|I|N|I|E|A|P|S|T|Y|R 620
 Db 1849 GAAATTAGCAACACCTTGCCCTTGAGAGGCGCCTCAATATGCGCAATAGAGATTAC 1908
 Qy 621 G|A|N|I|E|G|Y|A|P|V|A|I|E|U|L|E|U|S|E|R|I|E|G|I|N|I|E|U|S|E|R|S|E|R|A|A|I|A|I|S 640
 Db 1909 CAAAGATCCGGCGATGTCATCTGAAGAACATTCAGGGCATGAAGCACTGGCCGCTCAC 1968
 Qy 641 L|E|U|T|P|Y|H|I|S|E|R|G|U|A|L|E|U|G|U|A|L|E|U|G|U|N|E|N|G|Y|I|E|Y|S|E|R|S|E|R|A|G|Y 660
 Db 1969 CTGGAGACACACGAGCCTTGAGAGCCTTGAGAGATGGAATCAAGAGCTCCCGCGG 2028
 Qy 661 L|E|U|G|U|N|E|P|H|E|C|Y|A|G|A|P|P|H|E|G|U|N|E|U|N|Y|S|V|A|I|C|Y|S|T|Y|R|L|E|U|P|R|O|L|E|U|A|N|T|H|R 680
 Db 2029 CTGGAGAACTTCTCAGAGACTTGAGCTGCAAGAGGTGTCTCAACCCCTCAACACC 2088
 Qy 681 P|H|E|U|L|E|U|A|R|P|R|O|L|E|U|H|I|S|A|R|G|L|E|U|M|E|R|H|I|S|T|Y|R|S|G|I|N|V|A|L|L|E|U|G|U|A|R|G|L|E|U|C|Y|S 700
 Db 2089 TTCCCTCTGGGGCCCATCTGCACCGGCTCATGCACTACAGACAGTCTCGAGCGGCTGTGC 2148
 Qy 701 L|Y|E|H|I|S|I|P|R|O|S|E|R|H|I|S|A|A|P|P|H|E|A|R|G|A|P|C|Y|A|T|G|A|A|A|L|E|U|A|I|G|U|I|E 720
 Db 2149 AAACACACCCCGCGAGCCGACGCGCACTTCAGGGACCTGCGAGCCGCTTGGCAGAGATC 2208
 Qy 721 T|H|R|O|U|E|W|E|T|V|A|I|A|G|I|N|E|U|H|I|G|Y|T|H|E|T|I|E|Y|W|E|R|G|U|A|N|P|H|E|G|I|N|Y|L|E|U 740
 Db 2209 ACCGAGATGTGTGCACAGCTCCACGCTACGATGATCAAGATGGAATTTTCCAAAGCTG 2268
 Qy 741 H|I|S|I|U|L|E|U|Y|S|Y|A|P|L|E|U|I|E|G|Y|I|E|A|P|H|E|U|V|A|I|P|R|O|G|I|A|R|G|U|P|H|E 760
 Db 2269 CACGAATCTCAGAAAGATTTTATTTGGCATTCATCACTTGTGCTCCGGAAAGGAGTTTC 2328
 Qy 761 I|L|E|A|R|G|L|E|U|G|Y|S|E|R|S|E|R|Y|L|E|U|S|E|R|G|Y|L|E|U|G|I|N|A|R|G|M|E|R|P|H|E 780
 Db 2329 ATCCGCTCTGGGCACTCTCAGCAAGCTCTCGGGAAGGGGCTCCAGCAGCGCATGTCTTC 2388
 Qy 781 L|E|U|P|H|E|A|N|P|V|A|I|L|E|U|L|E|U|T|Y|R|T|H|S|E|R|A|R|G|Y|L|E|U|T|R|A|I|S|E|R|A|N|G|I|N|P|H|E|Y|S 800
 Db 2389 CTGTTCAACGACGTCCTGCTATACACGAGCGGGGCTGAGCGGCTCCAACTCACTTTAAA 2448
 Qy 801 V|A|I|E|G|Y|G|I|N|E|U|P|R|O|L|E|U|T|Y|R|G|Y|W|E|R|T|H|I|E|G|I|N|U|S|E|R|G|U|A|P|G|I|T|T|G|Y 820
 Db 2449 GTCCACGGGCACTCTCCGCTTATGCACTGACATTAAGAGACGAAAGCAAGTGGGG 2508
 Qy 821 V|A|P|R|O|H|I|S|C|Y|L|E|U|T|R|L|E|U|A|R|G|Y|G|I|N|A|R|G|I|N|S|E|R|I|E|I|V|A|I|A|I|S|E|R|S|E|R 840
 Db 2509 GTGCCCACTGCTGACCTCTCGGGGCAAGCGGCAATCATCATCTGCGCCAGTTCT 2568
 Qy 841 A|R|G|S|E|R|G|U|W|E|R|G|U|Y|T|T|R|V|A|I|G|U|A|P|I|E|G|I|N|E|C|A|I|I|E|A|P|L|E|U|A|I|G|U|Y|S 860
 Db 2569 CGGTCACAGATGGAAGTGGGTTCAGACATCCAGATGCGCATTTGACCTGGCCGAGAG 2628
 Qy 861 S|E|R|S|E|R|P|R|O|A|I|P|R|O|G|I|N|P|H|E|U|A|I|S|E|R|S|E|R|P|R|O|A|P|A|N|Y|S|E|R|P|R|O|A|P 880
 Db 2629 AGCGACGACCCCGCCTGATCTCTGCGCAGACGCCCCCTGCAACAACTCCCTGAT 2688
 Qy 881 G|U|A|I|A|T|H|R|A|I|A|A|P|G|I|N|U|S|E|R|G|U|A|P|A|P|L|E|U|S|E|R|A|I|S|E|R|A|R|G|I|T|H|S|E|R|U 900
 Db 2689 GAAGCCACCGCGGCTGACAGAGATCAAGAGATGACCTGACCGCTCGCCAGCATCCGCTG 2748
 Qy 901 G|U|A|R|G|I|N|A|P|R|O|H|I|S|A|R|G|Y|A|N|T|H|E|T|V|A|I|H|S|V|A|I|C|Y|T|R|P|H|I|S|A|R|A|N|T|H|R 920
 Db 2749 GAGGCGAGGCGCGCACCGCGCAACACATGTGTGACGCTGTGCAACCGCAACACC 2808
 Qy 921 S|E|R|V|A|I|S|E|R|E|T|V|A|I|S|E|R|S|E|R|I|E|I|A|I|A|G|U|A|N|G|I|N|U|S|E|R|G|Y|A|S|N|L|E|U 940
 Db 2809 AGGCTCTCCATGTGTGACTTACGATCGCACTGAGATCGATGTTGTCTGGAACCTGCTG 2868
 Qy 941 A|R|G|Y|S|P|H|E|Y|A|N|S|E|R|A|N|G|Y|T|T|R|G|I|N|Y|L|E|U|T|R|V|A|I|P|H|E|R|A|N|P|H|E|C|Y|S 960
 Db 2869 AGGAAATTCAAAACACGAAACGGGTGGCAAGAGCTGTGGGTGTTCACAAACTTCTGC 2928

Qy 961 L|E|U|P|H|E|P|H|E|T|Y|R|S|E|R|H|I|S|A|P|A|S|N|H|I|P|R|O|L|E|U|A|I|S|E|R|L|E|U|P|R|O|L|E|U|G|Y 980
 Db 2929 CTGTTCTTCTCAATATCACACGAGACAAATATATCCCTTGCACCTGCTTCTGCGC 2988
 Qy 981 T|Y|R|S|E|R|L|E|U|T|R|I|L|E|P|R|O|S|E|R|G|U|S|E|R|G|U|A|N|I|E|G|I|N|Y|S|A|P|Y|R|V|A|I|P|H|E|Y|S|E|U 1000
 Db 2989 TACTCGTCAACATCCCTCTGAGTCCGAGAACATCCAGAAAGCTACGTTCAAGCTG 3048
 Qy 1001 H|I|S|P|H|E|Y|S|E|R|H|I|S|V|A|I|T|Y|R|P|H|E|A|R|G|A|I|G|U|S|E|R|G|U|Y|R|T|H|R|P|H|E|G|U|A|R|G|T|P 1020
 Db 3049 CACTTCAAGTCCCGACGCTACTACTTACGAGCGGGAAGCAAGTACAGTTGAAAGGTGG 3108
 Qy 1021 M|E|R|G|U|V|A|I|I|E|A|R|S|E|R|A|I|T|H|S|E|R|S|E|R|A|I|S|E|R|P|R|O|H|I|S|V|A|L|L|E|U|S|E|R|H|I|S|Y|S 1040
 Db 3109 ATGGAAGTGAATCCGACAGTGCACAGCTGTGCTTGCAACCCCGATGTGACCAAA 3168
 Qy 1041 G|U|S|E|R|L|E|U|V|A|I|T|Y|R 1045
 Db 3169 GAGTCTTGTGTAT 3183
 RESULT 2
 ABE03370 ID ABE03370 standard; DNA; 3442 BP.
 AC ABE03370;
 DT 09-FEB-2006 (first entry)
 DE XX
 KW Acute myelogenous leukemia prognosis related DNA sequence SEQ ID NO: 31.
 KW dg; gene; acute myelogenous leukemia; prognosis; gene expression;
 KW biochip.
 OS Homo sapiens.
 XX JP200533987-A.
 PN 08-DEC-2005.
 PD 06-MAY-2005; 2005JP-00135284.
 PF 06-MAY-2004; 2004US-0568635P.
 PR (VERT-) VBRIDEX LLC.
 PA Lapont M;
 PI WPI; 2006-004067/01.
 DR
 XX
 PT Evaluating an acute myelogenous leukemia patient's prognosis, comprising
 PT detecting lower or higher expression level of gene recognized by probe
 PT set e.g. 202820-at and 206148-at, with respect to predetermined cut-off
 PT level.
 PS
 PS Example 5; SEQ ID NO 31; 60bp; Japanese.
 XX
 CC The present sequence is that of a human gene sequence which is claimed
 CC for use in evaluating the prognosis of patients suffering from acute
 CC myelogenous leukemia (AML) by analysis of the patients gene expression
 CC profile. The invention relates to a novel method for evaluating the
 CC prognosis of a patient with acute myelogenous leukemia by detecting a
 CC higher/lower expression level of genes encoding mRNA recognized by a
 CC probe set chosen from 19 nucleotide sequences (ABE03340 or ABE03344-
 CC ABE03361). Also, gene expressions (ABE03362-ABE03527) using a set of
 CC probes based on 167 sequences (ABE03362-ABE03527) using a biochip was
 CC performed to determine prognosis. Methods are also included for
 CC determining an AML patient's state, determining AML patient's treatment
 CC protocol, determining whether the patient will respond to the treatment,
 CC and producing an AML patient's prognosis report by analysis of the
 CC patients gene expression profile. The methods are useful in evaluating an
 CC AML patient's prognosis and for treating AML patient.

Sequence 3442 BP; 864 A; 952 C; 927 G; 699 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	0	Length:	3442
Score:	5463.00	Matches:	1045
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	15	Gaps:	0

US-09-555-342B-2 (1-1045) x AEE03370 (1-3442)

QY 1 MetGlyLeuIleGluGlnArgProThrProGlySerArgLeuGlyAlaProGluAsnSer 20
DB 49 ATGGAGAAATTAAGCAGAGAGGCGGACCCAGGATCAGACTGGGGGCGCCGGAAAAATTCG 108
QY 21 GlyIleSerThrLeuGluArgGlyGlnLysProProThrProSerGlyLysLeuVal 40
DB 109 GGGATCAGTACTTGGAAACGTGCAGAGACCGCCCCAACACCTTCAGGAAAACTCGTG 168
QY 41 SerIleLeuIleGlnMetLeuAspThrGlnGluAlaPheGluValProGlnArgAla 60
DB 169 TCATCAAAATCCAGATGCTGGATGACACCCAGAGGCAATTGAAGTTCCAAAGAGCT 228
QY 61 ProGlyLysValLeuLeuAspAlaValCysAsnHisLeuAsnLeuValGluGlyAspTyr 80
DB 229 CTGGGAGAGGTGCTGCTGGATGCAAGTTTGCACACCACTCAACCTCGTGGAGAGGTACTAT 288
QY 81 PheGlyLeuGlnPheProAspHisLysLysIleThrValIlePheLeuAspLeuLysPro 100
DB 289 TTTGGCTCGAGTTTCTGTGATCAAAAAAGATCAGGTGTGGCTGCGATCTCTAAAAACC 348
QY 101 IleValIleGlnIleAspArgProLysHisValValLysPheValLysPhePhe 120
DB 349 ATTTGTAACAGATTAGAGGCCAAGACAGGTGTGTATTAAGTTGTGTGAATTTCTTT 408
QY 121 ProProAspHisThrGlnLeuGlnGluGlnLysLeuThrArgTyrLeuPheAlaLeuGlnVal 140
DB 409 CCGGCTTACCAACACAACTCCAAAGAAAGTCCAAAGGTAACCTGTTCGGGCTGCAGGTG 468
QY 141 LysGlnAspLeuAlaGlnGlyArgLeuThrCysAsnAspThrSerAlaAlaLeuLeuIle 160
DB 469 AAGAGAGACTTGGCTCAAGCAGAGGTTGACGTATATACACACAGCGCACTCTTGATTT 528
QY 161 SerHisIleValGlnSerGluIleGlyLysPheAspGluAlaLeuAspArgGlnHisLeu 180
DB 529 TCACACATTTGGCATTTGAAGATTTGGGATTTTGATGAAGCTTGGACAGAGACACTTA 588
QY 181 AlaLysAsnLysTyrIleProGlnGlnAspAlaLeuGlnLysPheValIleGluPheHis 200
DB 589 GCAAAAAATTAATACATACCTCAGCAAGACGCACTTAAGGCAAAATCGTGGAAATTTTCA 648
QY 201 HisAsnHisIleGlyGlnThrProAlaGluSerAspPheGlnLeuLeuGlnIleAlaArg 220
DB 649 CATTAACCACTTGGACAAACACCAAGCAAGATTCAGATTTCCAGCTCCCTAAGATTGGCCGT 708
QY 221 ArgLeuGlnLeuSerTyrGlyIleArgLeuHisProAlaLysAspArgGlnIleTyrLysIle 240
DB 709 CGGCTAGAGATGATGGAATCCGGTTCAACCGGCCAAGACAGGAGGACAGAGATC 768
QY 241 AsnLeuAlaValAlaAsnThrGlyIleLeuValPheGlnGlyPheThrLysIleAsnAla 260
DB 769 AATCTGGCCGTTGGCAACAGGGAAATTCATAGTTTCAAGGTTTCACTAAGATCAATGCC 828
QY 261 PheAsnThrAlaLysValAlaGlyLysLeuSerPheLysArgLysArgPheLeuIleLysLeu 280
DB 829 TTCAACTGGGCGCAAGGTGGAGCTGAGCTTCAAGAGGAAAGCGCTTTTCATCAACAGCTC 888
QY 281 ArgProAspAlaAsnSerAlaTyrGlnAspThrLeuGlnPheLeuMetAlaSerArgAsp 300
DB 889 GGGCGCATGCGCAATAGTGGCTACAGAGATACCTTGAAGATTTCTGATGGCCAGTCCGGAGT 948
QY 301 PheCysLysSerPheThrLysIleCysValGlnHisAlaPhePheArgLeuPheGlu 320

DB 949 TTCTGCAAGTCTTGGAAAAATCTGTGTGAACATCATAGCTCTTTAGACTTTTGAA 1008
QY 321 GluProLysProLysProLysProValIlePheSerArgGlySerSerPheArgPheSer 340
DB 1009 GAAGCCCAAAACCAAAAGGCCCAAGCCGTCCTTTAAAGCCGGGGGTCAATTCATTCGATTCAGT 1068
QY 341 GlyArgThrGlnLysGlnValLeuAspTyrValLysGluGlyLysIleLysLysValGln 360
DB 1069 GGTGGATCTGAGAGAGGTTCTCGACTATGTTAAAGAAAGAGACATAAGAGGTGCAG 1128
QY 361 PheGluArgLysHisSerLysIleHisSerIleArgSerLeuAlaSerGlnProThrGlu 380
DB 1129 TTTGAAGAGACACAGCAAGATTCAATCTATCCGAGCCTTGCTTCAACAGCTTACAGAA 1188
QY 381 LeuAsnSerGluValLeuGlnGlnSerGlnGlnIleSerThrSerLeuThrPheGlyGly 400
DB 1189 CTGAATTCGAAATGCTGGAGAGTCTCAGCAGAGACCAAGCTTACATTTGAGAAAGGT 1248
QY 401 AlaGluSerProGlyLysGlnSerCysArgArgGlyLysGluProLysValSerAlaGly 420
DB 1249 GCCGAATCTTCAGAGGGGCGCAGAGCTCCGGCGAGAAAGAACCGAAGTTTCCGCCGG 1308
QY 421 GluProGlySerHisProSerProAlaProArgArgSerProAlaGlyAsnLysGlnAla 440
DB 1309 GAAGCGGGGTGCAACCGAGCCTCGCGCGAGAGAGAACCGCGGGTAAACAAGCAGGCG 1368
QY 441 AspGlyAlaAlaSerAlaProThrGlnGlnGlnGlnValValLysAspArgThrGln 460
DB 1369 GACGAGAGCGGCTCGCGGCCACAGAGAGAAAGAGAGTCTTAAAGATAGAGACCCAG 1428
QY 461 GlnSerLysProGlnProGlnProGlnProSerThrGlySerLeuThrGlySerProHisLeu 480
DB 1429 CAGAGTAACTCAGGCCCGCCGACAGCCACAGAGCTCCTGACTGGCACTTCACTT 1488
QY 481 SerGlnLeuSerValAsnSerGlnGlyValAlaProAlaAsnValThrLeuSerPro 500
DB 1489 TCCGAGCTGTGTGAATCGCAGAGGGAGGTGGCCCTGCGCAACCTGACTGTCTGCC 1548
QY 501 AsnLeuSerProAspThrLysGlnAlaSerProLeuIleSerProLeuAsnAspGln 520
DB 1549 AACCTAAGCCCCACACCAAGAGGCTCTCTCTTATACAGCCGCTGCTGAATGACAG 1608
QY 521 AlaCysProArgThrAspAspGluAspGluGlyArgArgLysArgPheProThrAspLys 540
DB 1609 GCTTGCCCCGAGCGACATAGAGATGAGGCGCGAGAAAGATTTCCCACTGATTA 1668
QY 541 AlaTyrPheIleAlaLysGluValSerThrThrGlnArgThrTyrLeuLysAspLeuGlu 560
DB 1669 GCGTACTCATAGCTAAGAGAGTGTACACCGAAGCGAATATCTGAAGAGATCTCGAA 1728
QY 561 ValIleThrSerThrPheGlnSerThrValSerLysGluAspAlaMetProGlnAlaLeu 580
DB 1729 GTTATACATTCGTGGTTTCAAGACACAGTAGAGAAAGAGACCCATGCGGAAAGCACTG 1788
QY 581 LysSerLeuIlePheProAspPheGlnProLeuHisLysPheHisThrAsnPheLeuLys 600
DB 1789 AAAAGTCTCATNTCCCGAATTTTGAACCTTTGCAAAATTTCAATCTAATTTTTCAG 1848
QY 601 GluIleGlnGlnArgLeuAlaLeuThrGlnGlyArgSerAsnAlaGlnIleArgAspTyr 620
DB 1849 GAATATTGACCAACGACTTCCCTGTGGGAAGCGCTCAAAATCCCAAAATCAAGATTAC 1908
QY 621 GlnArgIleGlyAspValMetLeuLysAsnIleGlnGlyMetLysHisLeuAlaHis 640
DB 1909 CAAAGATGCGGCATATCATGTGAAGAAATTCAGAGGACATGAAGCACTCGGGGCTCAC 1968
QY 641 LeuThrLysHisSerGlnAlaLeuGlnAlaLeuGluAsnGlyIleLysSerSerArgArg 660
DB 1969 CTGTGGAAGCAGACAGGCTTTGGAGGCTCTGGAAATAGGATCAAGAGCTCCGGCGG 2028
QY 661 LeuGluAsnPheCysArgAspPheGluLeuGlnLysValCysTyrLeuProLeuAsnThr 680

Db 2029 CTGGAGAACTTCTGACAGACTTTGAGCTGCAAGAAAGTGTGTACTACCGCTCAACACC 2088
 Qy PheLeuLeuAaGProLeuHiaArgLeuMetHiaGlyArgInValLeuGluArgLeuCys 700
 Db 681 TTCTCTCCGCGGCCCATGACCGGCTCATGCACTCAAGCAGAGCTCTGAGCGGCTGCGC 2148
 Qy 701 LysHiaHiaProSerHiaEaIaAaPheArgAaPheCysAaGAAaIaIaLeuAaIaGluIle 720
 Db 2149 AAACACCAACCCCGCCGAGCCAGCCGACCTTCAGAGACTCCGAGCCGCTTTGGCAGAGATC 2208
 Qy 721 ThGtMetValaIaGlnLeuHiaGlyThMetIleTybMetGluAaPheGlnIaIaLeu 740
 Db 2209 ACGGAGATGATGGCAGACTCCACGGTACGATGATCAAGATGAGAAATTCACAGAACTG 2268
 Qy 741 HiaGluLeuGlybAaPheLeuIleGlyIleAaPheLeuValaProGlyAArgGluIaPhe 760
 Db 2269 CACCAACCTCAAGAAAGATTTATGGCAATCTTGCGTTCCGCGAAGGAGATTC 2328
 Qy 761 IleArgLeuGlySerLeuSerIleLeuSerGlybGlyLeuGlnGlnIaArgMetPhePhe 780
 Db 2329 ATCCGCTCGGCGAGCTTCAGCAAGCTCTCGGAGAAAGGGAGCTTCAGCAGCGCATGTTCTTC 2388
 Qy 781 LeuPheAaAaPheValaLeuLeuTyThSerArgIleLeuThraIaSerAaGlnPheIaIa 800
 Db 2389 CTGTTCAACGACGCTCTGCTATACACGAGCCGGGGGCTGACGGGCTCCAAATCAATTTAA 2448
 Qy 801 ValHiaGlyGlnLeuProLeuTyGlyMetThraIleGluGluSerGluAaArgIuTrrGly 820
 Db 2449 GTCCACGGGAGCTCCCGCTTATGGCAATGACATGAGAGAGAGAGAGAGAGAGAGAGAG 2508
 Qy 821 ValProHiaCysbLeuThraLeuArgGlyGlnArgInSerIleIleValaIaIaSerSer 840
 Db 2509 GTGCCCCACTGCTGACCTCCGGGGCCAGCGGCACTCAATGATGCGCCGCAATTTCT 2568
 Qy 841 ArgSerGluMetGluTybTrrValaGluAaPileGlnMetAlaIaAaPheLeuAaGluIaIa 860
 Db 2569 CGGTCGAGAGATGAGAAAGTGGTGAAGACATCCAGATGGCCATTTGACCTGCGCGAGAAAG 2628
 Qy 861 SerSerSerProAlaProGluPheLeuAaSerSerProPolaAaPheIaIaSerProAaP 880
 Db 2629 AGCAGCAGCCCCCGCTGAAATCTTGACCGAGCCCGCTGCAACAACTCCCTGAT 2688
 Qy 881 GluIaIaThraIaAaAaPngInGluSerGluAaPheLeuSerAlaSerArgThSerLeu 900
 Db 2689 GAAGCCACCGGCTGACAGAACTCAAGAACTCAAGCTGCGCTGCGCAATCGCTG 2748
 Qy 901 GluArgGlnIaProHiaArgGlyAaHiaThMetValaIaCysTrrPheIaArgAaHiaTh 920
 Db 2749 GAGGCGCAGGCCCCGACCGCGCAACAAATGGTGACGTTGCTGGCACCGCAACACC 2808
 Qy 921 SerIaIaSerMetValaAaPheSerIleAlaValaGluAaGlnLeuSerGlyAaPheLeu 940
 Db 2809 AGCGCTTCCAGTGGAGCTTCAGCATCGCAGTGAAGATCAAGTTGCTGGAACCTGCTG 2868
 Qy 941 ArgbAaPheIaAaAaSerAaPngIyTrrGlnIaIaSerLeuTrrValaIaPheThAaPheCys 960
 Db 2869 AGGAATTCACAAACACAGCAGCGGTGGCAAAAGCTGTGGGTTCACAAACTTCTGC 2928
 Qy 961 LeuPhePheTybIaSerHiaGlnAaPheAaHiaProLeuAaIaSerLeuProLeuGly 980
 Db 2929 CTGTTCTTTCACAAATCAACACAGCAATCATCCCTTGGCCAGCTGCGCTCTGCGGC 2988
 Qy 981 TybSerLeuThraIleProSerGluSerGluAaGlnIleGlnIaAaPrrTrrValaIaPheIaIa 1000
 Db 2989 TACTCGCTCAACATCCCTCGAGTCCGAAGAACTCCAGAAAGACTCAAGTTCAACGCTG 3048
 Qy 1001 HiaPheIaSerHiaIaValTybTrrPheArgIaGlnSerGluTybTrrPheGlnAaArgTrr 1020
 Db 3049 CACTTCAAGTCCACGTTACTACTTCAAGGCGGAAAGCAGATCAAGCTTCGAAAGGTGG 3108
 Qy 1021 MetGluValaIleArgSerAlaThSerSerAlaSerArgProHiaIaIaLeuSerHiaIaIa 1040
 Db 3109 ATGGAAGTGAATCCGACAGTGCACACAGCTCTGCTCGGGAACCCACGTTGAGCCACAAA 3168

Qy 1041 GluSerLeuValaTyb 1045
 Db 3169 GAGTCTCTGTGTAT 3183
 RESULT 3
 AEB22714
 ID AEB22714 standard; DNA; 5028 BP.
 AC AEB22714;
 XX 22-SEP-2005 (first entry)
 DT
 XX
 DE Human colon cancer-associated gene SEQ ID NO:9.
 XX
 KW cancer; genetic marker; lymphoma; cytostatic; hematological disease;
 KW immune disorder; neoplasm; colon tumor; gastrointestinal disease;
 KW adenocarcinoma; carcinoma; screening; diagnostic; prognostic;
 KW pharmaceutical; ds; gene.
 KW
 OS Homo sapiens.
 XX
 PN WO2005064009-A1.
 XX
 PD 14-JUL-2005.
 XX
 PF 23-DEC-2004; 2004WO-DK000914.
 XX
 PR 27-DEC-2003; 2003DK-00001940.
 PR 24-JAN-2004; 2004DK-0000096.
 PR 07-APR-2004; 2004DK-0000586.
 PR 26-NOV-2004; 2004DK-00001843.
 XX
 PA (AROS-) AROS APPLIED BIOTECHNOLOGY APS.
 XX
 PI Ortoft TF, Jensen JL, Krushoffer M, Laiho P, Aaltonen IA;
 DR WPI; 2005-506661/51.
 XX
 PT Classifying the cancer in an individual having contracted cancer
 PS comprises determining the microsatellite status of the tumor.
 XX
 PS Claim 9; SEQ ID NO 9; 338pp; English.
 CC The invention relates to a method of classifying the cancer in an
 CC individual having contracted cancer, which comprises determining the
 CC microsatellite status of the tumor. The microsatellite status or the
 CC hereditary or sporadic nature of the cancer is determined by a method
 CC comprising: (a) obtaining a sample from the individual having contracted
 CC cancer, the sample comprising gene expression products, the presence
 CC and/or amount of which forms a pattern that is indicative of the
 CC microsatellite status or the hereditary or sporadic nature of the cancer;
 CC (b) determining the presence and/or amount of the gene expression
 CC products forming the pattern, and obtaining an indication of the
 CC microsatellite status or the hereditary or sporadic nature of the cancer
 CC in the individual based on step (b). At least one or two of the gene
 CC expression products forming a pattern used to determine the
 CC microsatellite status or the hereditary or sporadic nature of the cancer
 CC are selected individually from any of the 134 genes comprising 367-7850
 CC bp (AEB22706 to AEB22839), given in the specification. Also included are:
 CC a method for treating an individual; a method for reducing malignancy of
 CC a cell; a method for producing antibodies against an expression product
 CC of a cell from a biological tissue; pharmaceutical composition, for
 CC treating a classified cancer, comprising at least one antibody, at least
 CC one polypeptide, or at least one nucleic acid and/or probe; and an assay,
 CC for classification of cancer in an individual having contracted cancer,
 CC comprising at least one marker capable of determining the microsatellite
 CC status in a sample and at least one marker in a sample determining the
 CC prognostic marker, where the microsatellite status and the prognostic
 CC marker is determined simultaneously or sequentially. The method above is
 CC useful for producing an assay for classifying cancer in animal tissue.
 CC The peptide, gene, or probe is useful for preparing a pharmaceutical
 CC composition for the treatment of a cancer in animal tissue. The method is

CC useful for classifying the cancer in an individual having contracted
CC cancer, i.e. colon cancer, e.g. an adenocarcinoma, a carcinoma, a
CC teratoma, a sarcoma, and/or a lymphoma. The present sequence represents a
CC colon cancer-associated gene used in the method of the invention.

XX Sequence 5028 BP; 1288 A; 1387 C; 1276 G; 1077 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	0	Length:	5028
Score:	5463.00	Matches:	1045
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	14	Gaps:	0

US-09-555-342B-2 (1-1045) X ABB22714 (1-5028)

QY 1 MetGlyGluIleGluGlnArgProThrProGlySerArgLeuGluValAProGluAenSer 20
Db 337 ATGGAGAAATAGACAGAGGCGGACCCAGGATCAGACTGGGGCGCCGGAATTCG 396
QY 21 GlyIleSerThrLeuGlnArgGlyGlnArgProProThrProSerGlyValLeuVal 40
Db 397 GGGATCAGTACTTGAGACGTGACAGAGCGGCCCAACCTTCAGGAAATCTGTG 456
QY 41 SerIleValIleGlnMetLeuAerPserThrGlnGluValPheGluValProGlnArgAla 60
Db 457 TCCATCAAAATCCAGATGCTGAGTGAACCCAGAGGACATTTGAAGTTCCAAAGAGCT 516
QY 61 ProGlyValValLeuLeuAerAlaValCysAenHileuAenLeuValGluGlyAerTyr 80
Db 517 CTGGGAGAGGTGGCTGAGTGCAGTTGGCAACCACTCAACCTCGGAGAGGUGACTAT 576
QY 81 PheGluValLeuGluPheProAerPheHileValIleThrValTrieuAerLeuLeuPro 100
Db 577 TTGGCTCCAGTTCTGTGATCAAAAGATGACGGGTGGCTGAGATCTCTAAACCC 636
QY 101 IleValIleGlnIleArgArgProCysHileValValIlePheValIlePhePhe 120
Db 637 ATTGTGAACAGATTGAGAGGCCAAGCAGCTGTTGTAAGTTGTGTGAATTTCTT 696
QY 121 ProProAerPheHileThrGlnLeuGlnGluLeuThrArgTyrIlePheAlaLeuGlnVal 140
Db 697 CCGCTGACCAACACACAACTCCAAGAACTCAAGAGTACCTGTTCGGCTGCAGGTG 756
QY 141 LysGlnAerPheuAglngIValArgLeuThrCysAenPserThrSerAlaIleuLeuIle 160
Db 757 AAGCAGACTTGGCTCAAGGAGGTGACGTGTATGACACACAGCAGCTCTTGAT 816
QY 161 SerHileValIleGlnSerGluIleGlyAerPheAerPheAlaLeuAerArgIleHileu 180
Db 817 TCACACTTGTGCAATCTGAGATTGGGATTTTGATGATAGCCTTGACAGAGCACTTA 876
QY 181 AlaIleAerAerPheTyrIleProGlnGlnAerAlaLeuGluAerPheIleValIlePheHile 200
Db 877 GCAAAATTAATATCATATCTCAGCAAGACGCACTAAGAGCAAAATCGGGAATTTT 936
QY 201 HisAenHileIleGlnIleThrProAgluSerAerPheGlnLeuGlnIleAlaArg 220
Db 937 CATTAACCAATTTGACAAACACAGCAAGATCAATTTCCAGCTCTAGAGATTTGCCGT 996
QY 221 ArgLeuGluMetGlyArgIleArgLeuHileProHileAerAerArgIleGlyThrValIle 240
Db 997 CGGTGAGAGATGATGGAATCCGGTTGACCCGGCCAAAGACAGGAGGACAGAGATC 1056
QY 241 AsnLeuAlaValAlaAerThrGlyIleLeuValPheGlnGlyPheThrValIleAenAla 260
Db 1057 AATCTGGCCCTTGGCAACAGCGGAATTTCTAGTGTTCAGGGGTTTCACTAAGATCAATGCC 1116
QY 261 PheAerThrAlaValAerValArgValLeuSerPheValArgValPheLeuIleValLeu 280
Db 1117 TTCACTGGGCGAAGGTGGAGCTGAGCTTCAAGAGGAAGGCGCTTTCATCAACAGCTC 1176

QY 281 ArgProAerAlaAenSerAlaTyrGlnAerThrLeuGluPheLeuMetAlaSerArgAer 300
Db 1177 CGGCCAGATGCCAATATGTGCTATCCAGAGTACCTTGGAATTCCTGATGGCAGTCCGGAT 1236
QY 301 PheCysLeuSerPheThrPheIleCysValGluHileAlaPhePheArgLeuPheGlu 320
Db 1237 TTCTGCAAGTCTTTCGGAATCTGTGTGAACATCATATGCTTTTGAGACTTTTGA 1296
QY 321 GluProIleProIleProIleProValLeuPheSerArgGlySerSerPheArgPheSer 340
Db 1297 GAGCCAAACCAAGGCCCAAGCCCTCTCTTAAGCCGGGGGTCACTATTCGGTTCAGT 1356
QY 341 GlyArgThrGlnIleGlnValLeuAerTyrValIleGluGlyGlyHileValValGln 360
Db 1357 GGTGGACTCAGAAAGCAGGTTCTCGACTATGTAAAGAAAGAGGACATTAAGAGGTGCAG 1416
QY 361 PheGluArgValHileSerValIleHileSerIleArgSerLeuAlaSerGlnProThrGlu 380
Db 1417 TTGAAAGAGACAGCAAGATTCATCTATCCGAGGCTTGCTTACAGCTTACAGAA 1476
QY 381 LeuAenSerGluValLeuGlnSerGlnSerThrSerLeuThrPheGlyGluGly 400
Db 1477 CTGAATTCGGAAGTCTGAGAGCATCTCAGCAGAGACCAAGCTTATTTGAGAAAGT 1536
QY 401 AlaGluSerProGlyGlyGlnSerCysArgArgGlyLysGluProIleValSerAlaGly 420
Db 1537 GCCGAATCTCCAGAGGGCCAGAGCTCGCGCAGAAAGAACCGAAGGTTTCCGCCGG 1596
QY 421 GluProGlySerHileProSerProAlaProArgArgSerProAlaGlyAerPheGlnAla 440
Db 1597 GAGCCGGGTTCGACCCGAGCCCTGCGGAGGAGAAAGCCCGGGGTAAACAGCAGGCG 1656
QY 441 AspGlyAlaAlaSerAlaProThrGlnGluGluGluValValIleAerAerThrGln 460
Db 1657 GACGAGCCGCTCGGCCGCCACAGAGAAAGAGAGAGGTCTTAAAGATAGCAACCG 1716
QY 461 GlnSerLysProGlnProProGlnProSerThrGlySerLeuThrGlySerProHileu 480
Db 1717 CAGAGTAAACCTCAGCCCGCCAGCCACACAGGCTCCCTGACTGGCAGTCCCTCACT 1776
QY 481 SerGluLeuSerValAenSerGlnGlyValAlaProAlaAerAerPheThrLeuAerPro 500
Db 1777 TCCGAGCTGTGTGAATCTGACGGGGAGAGTGGCCCTGCAACGTGACCTGTCTCC 1836
QY 501 AsnLeuSerProAerPheThrIleGlnAlaSerProLeuIleSerProLeuAenAerGln 520
Db 1837 AACCTAGGCCGCAKCAACAGAGGCTCTCTCTTATCAGCCGCTGCTGATGACAG 1896
QY 521 AlaCysProArgThrAerAerGluAerGluArgValArgValArgPheProThrAerPhe 540
Db 1897 GCCTGCCCGCGAGCATGATGAGTGAAGGCGCGGAGAAAGATTTCCCACTGATAA 1956
QY 541 AlaTyrPheIleAlaIleArgValIleSerThrThrGlnArgThrTyrIleuAerPheGlu 560
Db 1957 GCGTACTTCAATGCTAAGAAAGTGTACACAGCAAGCAATATGTGAAGATCTCGAA 2016
QY 561 ValIleThrSerThrPheGlnSerThrValSerLysGluAerAlaMetProGluAlaLeu 580
Db 2017 GTTATCATCTTCGTGTTTCAAGACAGTGAAGAAAGAGCAAGCAGTGCAGAGCACTG 2076
QY 581 LysSerLeuIlePheProAerPheGluProLeuHileValPheHileThrAerPheLeu 600
Db 2077 AAAAGTCTATATTTCCGAATTTTGAACCTTTGCAAAATTTTCAATTTTTCAG 2136
QY 601 GluIleGluGlnArgLeuAlaLeuThrPheGluArgSerAerAlaGlnIleArgAerTyr 620
Db 2137 GAAATGACCAAGACTTCCCTGTGGAAAGCGCTCAAAATGCCAAATCAAGAGATTC 2196
QY 621 GlnArgIleGlyAerValMetLeuAerAenIleGlnGlyMetCysHileValAerAlaHile 640
Db 2197 CAAGAAATCGGCAATCTATGCTGAAGAACTTCAAGGAGATGAAGCACTGGCGGCTCAG 2256
QY 641 LeuThrPheHileSerGluAlaLeuGluAenGlyIleLysSerSerAerArg 660

Db	2257	CTGTGGAAAGCAACGAGGAGCCCTTGGAGGCCCTGGAGGAATGGAAATTCMAAGCTCCCGGCGG	2316
Oy	661	LeuGIuaenPheCyAArgAaspPheGIuIeuGIuIysValCyAryrLeuProLeuAanThr	680
Db	2317	CTGGAGAACTTCTGCAAGACCTTTAGCTGCGAGAAAGGTGTATTACTTACCGCTCAACACC	2376
Oy	681	PheLeuLeuAargProLeuNIaArgLeuMetHsryrIyrgIuaIValIeuGIuaArgLeuCyA	700
Db	2377	TTCTCTCTGCGGCCACTGCAACCGGCTGACTGACATCAAGACAGGTCTTGAGCGGGCTGGCG	2436
Oy	701	LyHsHsHsPProProSerHsHsAlAaAPheArgAaPcyAaGAlaAlaLeuAlaGluIle	720
Db	2437	AAACACCAACCGGCCGACGACCGCCCACTTCAGGGACTCCGACGCGCTTTGGCAGAGATC	2496
Oy	721	ThrgIuMetValAlaGIuLeuNIaGIyHrMetCileuMetGIuaenPheGIuIyLeu	740
Db	2497	ACGGAGATGTGGCAACAGCTCCACGGATACGATGATCAAGATGAGAAATTTCCAGAGCTG	2556
Oy	741	HsGIuLeuLyIyAaAPLeuIleGIyIleAaPaenLeuValIaProGIYAArgIuPhe	760
Db	2557	CACGAACCTCAAGAAAGATTGATTGGCATTTGACATCTTGGTGTCCGGAGAGGAGATTCC	2616
Oy	761	IleArgLeuGIySerLeuSerLyLeuSerGIyLyGIyLeuGIuNGIuaArgMetPhePhe	780
Db	2617	ATCCCTCTGGGCGAGCTCAGCAAGCTCTCGGGGAAAGGGGCTCCAGCAACGACATGTTCTTC	2676
Oy	781	LeuPheAaAaPValIleuLeuTyrrHsSerArgGIyLeuThrAlaSerAaNGIuPheIys	800
Db	2677	CTGTTCACACGACGTCTCTATACACGAGCGGGGGCTGACGGCTCCAAATCAGTTTAA	2736
Oy	801	ValHsGIyGIuIeupProLeuTyrgIyMetThrIleGIuGIuSerGIuAaPGIuTrrGIy	820
Db	2737	GTCCACGGGCAAGCTCCCGCTCTATGGCATGAAGATTGAGAGAGGAAAGACAGATGGGGG	2796
Oy	821	ValProHsCyALeuThrLeuAargGIyGIuaArgIuSerIleIleValAlaAaSerSer	840
Db	2797	GTGCCCCACTGCTACCTCCCGGGGGCGAGCGGCACTCATCATCTGTGGCGGCAATTCCT	2856
Oy	841	ArgSerGIuMetGIuLyrrTrValGIuAaPrrIleGIuMetAlaIleAaPLeuAlaGIuIys	860
Db	2857	CGGTCCGAGATGAGAAATGGGTGTGAAGACATCCAGATGGCCATTGACTTGCGGAGAGAG	2916
Oy	861	SerSerSerProAlaProGIuPheLeuAlaSerSerProAaPaenLySerProAaP	880
Db	2917	AGCAGCAGCCCCCGCCCTGAGTCTCTGGCCACAGACCCCCCTGACAAACATCCCTCAT	2976
Oy	881	GIuaIaThrAlaIaAaPrrGIuSerGIuAaPLeuSerAlaSerAaArgThrSerIeu	900
Db	2977	GAAAGCAACGGCGCTGACCAAGAGTCAGAGAGTGACTGACGCGCTCCGCAATCGCTG	3036
Oy	901	GIuAargIuaIaProHsAaArgLyAaenHrMetValHsValCyAtrPrHsAaArgAanThr	920
Db	3037	GAGCGCCAGGGCCCCCGCACCGCGGCAACAAATGGCGACGTTGTCTGGCACCGCAACCC	3096
Oy	921	SerValaSerMetValaAaPheSerIleAlaValGIuaenGIuIeupSerGIyAaenLeu	940
Db	3097	AGCGTCTCAAGTGGACTTCAGCATCCCAATGGAAATCAAGTTGTCTGGAAACCTGCTG	3156
Oy	941	ArgLyPheLyAaenSerAaNGIyTrrGIuIyLeuTrrValIaPheThrAaenPheCyA	960
Db	3157	AGGAATTCAAAAACAGCAACGAGGAGGCAAGAGCTGGGGGTGGTTCACAAACTTCTGC	3216
Oy	961	LeuPhePheTyrrLyAaSerHsGIuAaPaenHsPProLeuAlaSerLeuProLeuLeuGIy	980
Db	3217	CTGTCTTCTTCAAAATCAACACGAGACATATCCCTTGCACGCTGCTCTCTCTCGGC	3276
Oy	981	TyrrSerLeuThrIlePProSerGIuSerGIuAaenIleGIuIyAaPrrTyrrValPheLyLeu	1000
Db	3277	TACTGCTCACCATCTCTCTGAGTCCGAGAACATCCGAAAGACTACGTCTTCAAGCTG	3336
Oy	1001	HsPheLySerHsValTyrrTyrrPheAaArgAlaGIuSerGIuTyrrHsPheGIuAargTrr	1020

Db	3337	CAC TTCGAAGTCCCAAGCTCTACTTACTTCAAGGCGGAAAGCAGTACAGTTCGAAAGGTGC	3396
Qy	1021	MetGluValIleArgSerAlaThrSerSerAlaSerArgProHisValIleuSerHisIlys	1040
Db	3337	ATGGAAGTGAATCCGACAGTCCACACAGCTCTGCTCGGACCCCAAGTGTGAGTCACAA	3456
Qy	1041	GluSerIeuValIyIyr	1045
Db	3457	GAGTCTCTGTGTAT	3471
RESULT 4			
AAA08582	ID	AAA08582 standard; DNA; 4687 BP.	
XX	AC	AAA08582;	
XX	XX	19-JUL-2000 (first entry)	
DT	XX		
DE	XX	Human cytoskeleton associated protein 2 (CYSKP-2) coding sequence.	
XX	XX	Cytoskeleton associated protein; CYSKP-2; cancer; proliferative;	
KM	XX	autoimmunity; inflammatory; vesicle trafficking; neurological;	
KM	XX	cardiovascular; cell motility; reproductive; muscle disorder; ss.	
XX	OS	Homo sapiens.	
XX	XX		
FH	XX	Key	Location/Qualifiers
FT	XX	CDS	90..3327
FT	XX		/*tag= a
FT	XX		/product= "CYSKP-2"
PN	XX	WO200017355-A2.	
PD	XX	30-MAR-2000.	
PF	XX	17-SEP-1999; 99WO-US021565.	
PR	XX	18-SEP-1998; 98US-0172226P.	
PR	XX	27-APR-1999; 99US-011321P.	
PA	XX	(INCY-) INCYTE PHARM INC.	
PI	XX	Lal P, Tang YT, Yue H, Hillman JF, Bandman O, Corley NC;	
PI	XX	Guessler KJ, Patterson C, Azimzai Y, Baughn MR;	
XX	XX	WPI; 2000-283582/24.	
DR	XX	P-PSDB; AAY91947.	
PT	XX	Human cytoskeleton associated proteins, used to treat cell proliferative,	
PT	XX	autoimmune/inflammatory, vesicle trafficking, neurological, cell	
PT	XX	motility, reproductive and muscle disorders.	
PS	XX	Claim 9; Page 101-102; 113pp; English.	
XX	XX	AAA08581-96 encode human cytoskeleton associated proteins 1 to 16 (CYSKP-	
CC	XX	1 to CYSKP-16) respectively. The sequences can be used to treat and	
CC	XX	diagnose cancer and cell proliferative, autoimmune/inflammatory, vesicle	
CC	XX	trafficking, neurological, cardiovascular, cell motility, reproductive	
CC	XX	and muscle disorders. Pharmaceutical compositions containing CYSKP-2 to	
CC	XX	CYSKP-16 can be used to treat or prevent disorders associated with	
CC	XX	decreased expression or activity of CYSKP (claimed), for example,	
CC	XX	autohemoclasiosis, cirrhosis, hepatitis, myelofibrosis, psoriasis, asthma,	
CC	XX	acquired immunodeficiency syndrome (AIDS), Crohn's disease, diabetes	
CC	XX	mellitus, Goodpasture's syndrome, osteoarthritis, scleroderma and trauma.	
CC	XX	CYSKP antagonists can be used to treat or prevent a disorder associated	
CC	XX	with increased expression or activity of CYSKP (claimed)	
XX	XX		
SO	XX	Sequence 4687 BP; 1179 A; 1261 C; 1196 G; 1051 T; 0 U; 0 Other;	
Alignment Scores:			
Pred. No.:	0	Length:	4687
Score:	5459.00	Matches:	1044

Percent Similarity: 99.9% Conservative: 0
Best Local Similarity: 99.9% Mismatches: 1
Query Match: 99.9% Indels: 0
DB: 3 Gaps: 0
US-09-555-342b-2 (1-1045) x AA08582 (1-4687)

QY 1 MetGlyIleuIleugInaRProthPrGlySerargLeuGlyAlaProGluAaSer 20
Db 90 ATGGAGAAATAGACAGAGCCGACCCAGATACAGACTGGGGGCCCCGGAATAATCG 149
QY 21 GlyIleSerThrIleuGlyArgGlyGlnIleProProProProThrProSerGlyIleVal 40
Db 150 GGATCATGATCACTTGGAACCTGGACAGAACCCGCCCAACCTTCAGAGAAAATCTGTG 209
QY 41 SerIleValIleGlnMetIleuAaPheThrGlnGlnAlaPheGlnValProGlnaGlyAla 60
Db 210 TCATCAAAAATCCAGAGCTGTGATGACACCCAGAGGCAATTTGAAGTTCCAAAGAGCT 269
QY 61 ProGlyIleValIleuAaPheAlaValCyAaSerIleuAaSerValGlnGlyAaPheTyr 80
Db 270 CTGGAGAAAGTCTGCTGATGATGCACTTTCACCACTCACTCGTGGAAAGTGAATAT 329
QY 81 PheGlyIleuGlnPheProAaPheIleValSerIleThrValIlePheAaPheIleuIlePro 100
Db 330 TTGGCTCGAGTTTCTGATCACAAAAAGATCAAGGTGTGGCTGATCTCTAAACC 389
QY 101 IleValIleGlnIleAaArgProIleValIleValIlePheValIleValPhePhe 120
Db 390 ATGTGAAACAGATTAGAGGACCAAGACGTTGTGTATGATTTGTGTGAATTTCTT 449
QY 121 ProProAaPheIleThrGlnIleuGlnGlnIleuIleThrArgTyrIleuPheAlaIleuGlnVal 140
Db 450 CCGGCTGACCAACACAACTCCAAGAGAACTCAAGATACCTGTTCCGGCTGACAGGTG 509
QY 141 LysGlnAaPheValaGlnGlyArgIleuThrCyAaAaPheThrSerAlaAlaIleuIle 160
Db 510 AACGAGACATTTGGCTCAAGGACAGGTGTGACGTATATACACAGCGACGCTCTTGATT 569
QY 161 SerHisIleValaGlnIleuGlnIleuIleGlyAaPheAaPheGlnAlaIleuAaPheGlnIleu 180
Db 570 TCAACATTTGCAATCTGAATTTGGGATTTTGATGAAGCTTTGACAGAGACACTTA 629
QY 181 AlaIleAaPheIleuIleProGlnIleuAaPheAlaIleuGlnAaPheIleValaGlnPheHis 200
Db 630 GCAGAAAATTAATACATCTCAAGCAAGACGCACTTAAGACAAATCGTGAATTTTCA 689
QY 201 HisAaPheIleGlyGlnIleProAlaGlnIleuAaPheGlnIleuIleAlaArg 220
Db 690 CATTAACCACTTGACAAACACCGAGCAATTCAGATTTCCAGCTCCCTAGAGATTTGCCCT 749
QY 221 ArgIleuGlnIleuIleArgIleuIleArgIleuHisProAlaIleAaPheArgGlnIleuIleArg 240
Db 750 CGGCTAGAGATGATGGAATCCGGTTGACCCGCGCAAGACAGAGAGGACAGAGATC 809
QY 241 AaPheIleValaIleAaPheThrGlyIleuValaPheGlnGlyPheThrIleValaAaPheAla 260
Db 810 AATCTGGCCCTTGCAACAGGAAATCTAGTGTTCAGGGTTTCACTTAAGATCAATGCC 869
QY 261 PheAaPheThrAlaIleValaIleGlyIleuSerPheIleAaPheArgIleuIleValIleu 280
Db 870 TTCAACCTGGCCCAAGGTCGAGCTGACCTTCAGAGGAAAGCGCTTTTTCATCAACGCTC 929
QY 281 ArgProAaPheIleAaPheSerAlaIleArgIleuAaPheThrIleuGlnPheIleuMetAlaSerArg 300
Db 930 CGGCGAATGCAATAGTGTGATCAGAGATACCTTGAAATTCCTGATGGCCAGTCCGGAT 989
QY 301 PheCyAaPheSerPheThrIleuCyAaValaGlnIleHisAlaPhePheArgIleuPheGln 320
Db 990 TTCTGCAAGTCTTCTGAAATCTGTGTGAACATCATGCTTCTTAAAGCTTTTGA 1049
QY 321 GlnProIleuArgProIleuArgProValIleuPheSerArgIleuSerPheAaPheSer 340

Db 1050 GAGCCCAAAACAAAGCCCAAGCCGCTCTTTAGCCGGGGCTCATCATTTCCGTTCACT 1109
QY 341 GlyArgThrGlnIleuGlnValIleuAaPheTyrValIleGlnGlyIleValIleValaGln 360
Db 1110 GCTCGAATCTCAAGAGAGGTTCTCGACTATGTTAAAGAGAGACATTAAGAGGTCCAG 1169
QY 361 PheIleuArgIleHisSerIleHisSerIleArgSerIleuAlaSerGlnProThrGln 380
Db 1170 TTGAAGAGAGACAGCAAGATTCATTCATCCGAGCTTGCTTCAACAGCTTAAGAA 1229
QY 381 LeuAaPheSerGlnValIleuGlnIleuGlnIleuGlnIleuThrSerIleuPheGlnIleuGln 400
Db 1230 CTGAATTCGAAAGTGTGAGAGCTTCAGACAGAGCAAGCTTACATTTGAGAAAGGT 1289
QY 401 AlaGlnSerProGlyIleGlnIleuSerCyAaArgGlyIleGlnProIleValSerAlaGly 420
Db 1290 GCCGATCTCCAGGGGGCCAGAGCTCCGCGAGAGAAAGAACCGAAAGTTCCGCCGG 1349
QY 421 GlnProGlySerHisPheProAlaProArgArgSerProAlaIleValIleGlnAla 440
Db 1350 GAGCCGGGGTCCGACCCGAGCTCCGCGCGAGAGAGAACCCCGGGGTAAACAGAGCG 1409
QY 441 AaPheValaAlaSerAlaProThrGlnIleuGlnIleuValaIleValAaPheArgThrGln 460
Db 1410 GACGAGCGGCTCCGCGGCCACGAGAGAGAGAGAGAGAGTGTTAAGATTAAGACCCAG 1469
QY 461 GlnSerIleProGlnIleuProGlnIleuProSerThrGlySerIleuThrGlySerProHisIleu 480
Db 1470 CAGAGTAACTCAAGCCCGCCGACGACCAAGCAAGAGCTCCCTGATCGGAGTCTCACTT 1529
QY 481 SerGlnIleuSerValaAaPheGlnIleuValaAlaProAlaAaPheValIleuIleuSerPro 500
Db 1530 TCCGAGCTGTGTGAACTCGACAGGGGGAGTGGCCCTCGCAACCTGATCTGTCTCC 1589
QY 501 AaPheIleuSerProAaPheThrIleuGlnAlaSerProIleuIleSerProIleuAaPheGln 520
Db 1590 AACCTAGCCCGACACCAAGAGGCTCTCTCTTATCAAGCCGCTGATTAAGACCAAG 1649
QY 521 AlaCyAaPheArgThrAaPheGlnIleuGlnIleuArgIleuAaPheArgPheProThrAaPhe 540
Db 1650 GCTGCCCCCGAGCGAGCATAGAGATGAAGGGCGGAGAGAAAGATTTCCCACTGATTA 1709
QY 541 AlaTyrPheIleAlaIleGlnValaSerThrThrGlnArgThrTyrIleuIleAaPheIleuGln 560
Db 1710 GCGTACTTATAGCTTAAGAGAGTGTCTACACCGAGCAACATATCTGAAGATCTCGAA 1769
QY 561 ValIleThrSerThrPheGlnIleuThrValaSerIleuGlnAaPheIleuProGlnIleuIleu 580
Db 1770 GTTATCACTTCGTGTGATTAAGACAGAGTGAAGAAAGAGACGCAATGCCGGAAGCACTG 1829
QY 581 LysSerIleuIlePheProAaPheGlnProIleuHisIleuPheHisThrAaPheIleuIle 600
Db 1830 AAAAGTCTCATATTCGGAATTTTGAACCTTTGCAAAATTTCAATTAATTTTCAAG 1889
QY 601 GlnIleGlnGlnArgIleuAlaIleuTrpGlnIleuArgSerAaPheAlaGlnIleuArgAaPheTyr 620
Db 1890 GAAATTGACCAACGACTTCCCTGTGGGAAGGCGCTCAAAATGCCCCAAATCAAGAGATTAC 1949
QY 621 GlnArgIleGlyAaPheValIleuIleuIleuAaPheIleGlnIleuIleHisIleuAlaHis 640
Db 1950 CAAGAGATCGGCATATCATGTGAAGAAATTCAGAGGCACTTGAAAGCACTGGGGGCTCAC 2009
QY 641 LeuTrpIleHisSerGlnAlaIleuGlnAlaIleuGlnAaPheIleuIleuIleuSerArgArg 660
Db 2010 CTGTGAAGCAACAGAGGCTTTGAGGGCTGTGAAATAGATCAAGAGCTCCGGCGG 2069
QY 661 LeuGlnAaPheCyAaArgAaPheGlnIleuGlnIleuValaCyAaTyrIleuProIleuAaPheThr 680
Db 2070 CTGAGAACTTCTGCAAGACTTGTGAGTGAAGAGTGTGTAACTTACCGCTCAACACC 2129
QY 681 PheIleuIleuArgProIleuHisArgIleuMetHisTyrIleuGlnValaIleuGlnArgIleuCyA 700
Db 2130 TTCTCTCGGGCCACTGACCGGCTCATGCACTAAGAGAGGTCTTGAGCGGCTGTGC 2189

```

QY 701 LysHSHAPROSESHIALAASPPhenArgpCySArgAlaAlaLeuAlaGluLe 720
DB 2190 AAAAAACCCCGGAGCCAGCCGACTTCAGAGCTGGCCAGCGCTTGGCAGAGTTC 2249
QY 721 ThrGluMetValaAgInLeuHISGlyTherMetIleuysMetGluasnPhenGlnLeu 740
DB 2250 ACGGAGATGGTGCGCAGCTCCACGGTACGATGATCAAGATGGAGAAATTTCCAGAGGTG 2309
QY 741 HisGluLeuYblybAapLeuIleGlyIleAapbAneUValValProGlyArgGluPhe 760
DB 2210 CACCAACTCAGAAAGATTGATGGCATGACATCTTGCTGGCGGAAGGAGATTCC 2369
QY 761 IleArgLeuGlySerIleuSerIleuSerGlyYblyGlyLeuGlnGlnApxMetPhePhe 780
DB 2370 ATCCGCTGGGCGACCTTCAGAGCTCTCGGGGAAAGGGGCTCCAGCAGCGCATGTCTTC 2429
QY 781 LeuPheAapbAneValLeuLeuTyThrSerArgGlyLeuThrAlaSerAapbInPheYb 800
DB 2430 CTGTTCAACGACGCTCTGCTATACAGAGCGGGGGCTGAGCGCTCCAAATCAATTAA 2489
QY 801 ValHisGlyGlnLeuProLeuTyGlyMetThrIleGluGluSerGluAapGluTrpGly 820
DB 2490 GTCCACGGGCGAGCTCCGCTCATGGCATGACGATTGAGAGACGAAAGCAGATGGGGG 2549
QY 821 ValProHisCybLeuThrLeuArgGlyGlnAapGlnSerIleIleValAlaAlaSerSer 840
DB 2550 GTGCGCCACCTGGCTGAGCCCTCGGGGCGCAGCGCATCATCATCGTGGCCGCAATCT 2609
QY 841 ArgSerGluMetGluYblyTrpValGluAapIleGlnMetAlaIleAapLeuAlaGluYb 860
DB 2610 CGGTCCGAGATGGAGAGTGGGTGAGGACATCCAGATGGCCATTGACCTGGCGGAGAG 2669
QY 861 SerSerSerProAlaProGluPheLeuAlaSerSerProPheAapbAneYbSerProA 880
DB 2670 AGCAGCAGCCCCCGCTGAGATTCTCGGCGCAGACCCCCCTGCAACAAGTCCCTGAT 2729
QY 881 GluAlaThrAlaAlaAapbGlnGluSerGluAapbAneUValaSerArgThSerLeu 900
DB 2730 GAAAGCCACCGGGGCTGAGCAGAGATCAGAGATACCTGAGCGCTCGCGCAGATCCGTG 2789
QY 901 GluArgGlnAlaProHisArgGlyAapbTherMetValHisValCybTrpHisArgAapb 920
DB 2790 GAGGCGCAGCGCCCGCAGCCGGGCAACCAATGGTGACGCTGCTGGCAGCCGCAACC 2849
QY 921 SerValSerMetValaAapbSerIleAlaValGluAapbGlnLeuSerGlyAapbLeu 940
DB 2850 AGCGTCTCATGTGGACTTCAGCATCGCAGTGGAGATCAGTTGTCTGGAACCTGCTG 2909
QY 941 ArgGlybPheYbAapbSerAapbGlyTrpGlnYbLeuTrpValValPheThrAapbPhe 960
DB 2910 AGGAAATTCAAAAACACCAAGGGTGAGAAAGTGGGGGTGCTTCAACAACCTTCTGC 2969
QY 961 LeuPhePheTyblybSerHisGlnAapbAneUValaSerProLeuAlaSerLeuGluYb 980
DB 2970 CTGTTCTTTCACAAATACACACAGCAATATCCCTTGGCAGCTGCTCTCTGGGC 3029
QY 981 TybSerLeuThrIleProSerGluSerGluAapbIleGlnYbAapTybValPheYbLeu 1000
DB 3030 TACTCGCTCACATCCCTCTGAGTCCGAGAACATCCGAAAGACTAGTCTCAAGCTG 3089
QY 1001 HisPheYbSerHisValTybTybPheArgAlaGluSerGlyTybThrPheGlnAapTrp 1020
DB 3090 CACTTCAGATCCACGCTCTACTACTTCAAGGGCGGAAAGCGAGTTCGAAAGGGTG 3149
QY 1021 MetGluValIleAapSerAlaThSerSerAlaSerArgProHisValIleuSerHisGly 1040
DB 3150 ATGGAAGTGAATCCGACAGGCCACAGCTCTGCTCGGACCCACGTTTGAATCACA 3209
QY 1041 GluSerLeuValTyb 1045
DB 3210 GAGTCTCTGTGTAT 3224

```

```

RESULT 5
ID AAS64760
AC AAS64760;
DT 13-FEB-2002 (first entry)
DE DNA encoding novel human diagnostic protein #564.
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder; ss.
OS Homo sapiens.
XX MO200175067-A2.
XX 11-OCT-2001.
XX 30-MAR-2001; 2001MO-US008631.
XX
XX 31-MAR-2000; 2000US-00540217.
XX 23-AUG-2000; 2000US-00649167.
XX (HYSE-) HYSEQ INC.
XX Dmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
XX P-PSDB; ABG00573.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX Claim 1; SEQ ID NO 564; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptides and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
XX ftp.wipo.int/pub/publshd_pct_sequences
XX
XX Sequence 3187 BP; 810 A; 874 C; 876 G; 627 T; 0 U; 0 Other;
SQ
Alignment Scores:
Pred. No.: 0 Length: 3187
Score: 5431.00 Matches: 1042
Percent Similarity: 99.6% Conservative: 0
Best Local Similarity: 99.6% Mismatches: 3
Query Match: 99.4% Indels: 1
DB: 5 Gaps: 0
US-09-555-342B-2 (1-1045) x AAS64760 (1-3187)
QY 1 MetGlyGluIleGluGlnArgProThrProGlySerArgLeuGlnAlaProGluAapbSer 20

```


49 ATGGAGAAATAGAGCAGAGCCGACCCCGAGATCAGACTGGGGGCCCCGAAAAATTCG 108
Qy 21 G1Y1LSeSerThLeuG1uArG1yG1uYsProProthThProSeSerG1yLysLeuVal 40
Db 109 GGAGTTCAGTACCTTGGAACCTTGACAGAGCCGCCCAACCTTCAGAAAACTCTGTG 168
Qy 41 Ser1eLys1eG1mELeuAspAspThrG1nG1uA1aPheG1uVal1ProG1nArG1a 60
Db 169 TCCATCAAAATCCAGATGCTGGATGACACCCAGAGGACATTTGAAgTTTCCAAAAAGGCT 228
Qy 61 ProG1yLysVal1LeuLeuAspAlaVal1CyAspN1s1eLeuN1eVal1G1uG1yAspTyr 80
Db 229 CCTGGGAAGGTGCTGCTGAGATGCACTTTCACACCTCAACCTCGTGGAAAGGTGACTAT 288
Qy 81 PheG1yLeuG1uPheProAspN1s1yAsp1LeThrVal1TrpLeuAspLeuLeuYsPro 100
Db 289 TTTGGCTTCAGTTCTCTGATCACAATAAGATCAGGTGTGGCTGGATCTCTAAACCC 348
Qy 101 1LeVal1yG1n1LeArGArGProLYsN1sVal1Val1yAspPheVal1yAspPhe 120
Db 349 ATTGTGAACAGATTAGAGGCCAAAGCACGTTGTTGTAAgTTTGTGTGAATTTCTTT 408
Qy 121 ProProAspN1s1ThrG1nLeuG1nG1uLeuThArG1yTr1eupheAla1eug1nVal 140
Db 409 CCCCCTGACCAACACAACTCCAAAGAACTCAAGAGTACCTGTTCCGCTGCAAGTGC 468
Qy 141 LysG1aAspLeu1aG1nG1yArG1eThrCyAspAspThrSer1aAla1eLeu1e 160
Db 469 AAGCAGACTTGGCTCAAGCAGGTGACGTGTATACACAGCGACGCTCTTGATTT 528
Qy 161 SerN1s1eVal1G1nSerG1u1eG1yAspPheAspG1uA1aLeuAspArG1uN1s1e 180
Db 529 TCACACATTTGCAATCTGAGATTGGGATTTTGATGAAGCTTTGACAGAGCACTTA 588
Qy 181 A1aLysAspN1s1yTr1eProG1nG1nAspAla1eug1uAspLys1eVal1G1uPheN1s 200
Db 589 GCAAAAATAATAACATACCTCAGCAAGACGCACTAAGAGCAAAATCGGGAATTTCCAC 648
Qy 201 H1aAspN1s1eG1yG1nThrProAlaG1uSerAspPheG1nLeuG1u1eAlaArG 220
Db 649 CATTAACCATTTGGACAAACACAGCAGAAATCAGATTTCCAGCTCTAGAGATTGCCCT 708
Qy 221 ArG1eug1uMetYrG1y1eArG1eN1s1eProAlaLysAspArG1uG1yTh1y1e 240
Db 709 CGGCTAGAGATGATGGAATCCGGTTCACCCGGCCAAAGCAGGGAAGCAGAGATC 768
Qy 241 AsnLeuN1aVal1aAspThrG1y1eLeuVal1PheG1nG1yPheTh1y1eAspAla 260
Db 769 AATCTGGCCCTTGCAACAGGGAAATTTCTAGTGTTCAGGGGTTTCATTAAGATCAATGCC 828
Qy 261 PheAspThrAlaLysVal1ArG1y1eLysSerPheLysArG1yAspPheLys1eLysLeu 280
Db 829 TTCAACTGGGCCAAAGGCGGAAGCTAGCTTCAGAGGAAGGCTTTTCTCA1CAACATC 888
Qy 281 ArGProAspAlaAspSerAlaYrG1nAspThr1eug1uPhe1eMetAlaSerArGAsp 300
Db 889 CGGCAAGATCCAAATAGTGGCTACAGAGATACCTTGAATTCCTGAATGCCAGTCCGGAT 948
Qy 301 PheCyLysSerPheTh1y1eCyAspVal1G1uN1s1eAlaPhePheArG1eupheG1u 320
Db 949 TTTGTGAAGCTCTTGGAATAATCTGTGTGAACATATGCTCTTTAACTTTTGA 1008
Qy 321 GluProLYsProLYsProLYsProLYsPheSerArG1ySerSerPheArGpPheSer 340
Db 1009 GAGCCCAACCAAGCCCAAGCCGCTCTTTAGCCGGGGGTCATATATTCGGTTCAGT 1068
Qy 341 G1yArGThrG1nYsG1nVal1LeuAspTyrVal1yG1uG1yG1yN1s1yLysVal1G1n 360
Db 1069 GGTCCGAGCTCAGAAAGAGGTTCTCGACTATGTTAAAGAGAGAGACATTAAGAGGTCCAG 1128
Qy 361 PheG1uArG1yN1s1eSer1y1eH1sSer1LeArGSer1eUAlaSerG1nProTh1yG1u 380

Db 1129 TTTGAAAGGAAGCAGCAAGATTCATTCGAGCTTGCTTCACAGCTCAAGAA 1188
Qy 381 LeuAspSerG1uVal1eug1uG1nSerG1nG1nSerTh1SerLeuThrPheG1yG1u-G1 400
Db 1189 CTGAATTCGGAAGTGTGAGAGCATGCTCAGCAGAGACACAGCTTATCTTTGGAGAGAGG 1248
Qy 400 yAlaG1uSerProG1yG1yG1nSerCyArGArG1yLysG1uProLYsVal1SerAlaG1 420
Db 1249 TGCAGATTTTCCAGGGGGCAGAGCTGCGGCAGAGAAAGAAACCAAGGTTTCCGCCGG 1308
Qy 420 yG1uProG1ySerN1s1eProSerProAlaProArGAspSerProAlaG1yAspN1yG1n1 440
Db 1309 GGAAGCCGGGGGTGCACCCAGAGCTGCGCGGAAGAAAGCCCGCGGTAAACAAGAGGC 1368
Qy 440 aAspG1yAlaAlaSerAlaProThrG1uG1uG1uG1uVal1yAspArGThrG1 460
Db 1369 GAGCGAGCCGCTCCGAGACCCAGAGGAAGAGAGAGGTGTTAAAGATAGAACCA 1428
Qy 460 nG1nSerLYsProG1nProProG1nProSerThrG1ySerLeuTh1ySerProN1s1e 480
Db 1429 GCAGAGTAAACCTCAAGCTCCGACGCCAAGCAGAGCTCCTGACTGCGAGTCCACCT 1488
Qy 480 uSerG1uSerVal1aAspSerG1nG1yVal1a1aProAlaAspVal1Thr1euphePr 500
Db 1489 TTTCCAGCTGTCTGTAACTCGAGGGGAGTGTGCCCTTGCCAAAGCTGACCTTGCTCC 1548
Qy 500 oAsnLeuSerProAspThr1yG1n1a1eSerProLeu1eSerProLeuLeuAspG1 520
Db 1549 CAACCTGAGCCCGCAACACAGAGGCTCTCCCTTGATCAAGCCCTGTGATAGACCA 1608
Qy 520 nAlaCyAspArGThrAspAspG1uAspG1uG1yArGArG1yLysArGpPheProTh1yAspLys 540
Db 1609 GGCCTGCCCCCGGACGAGATGAGGATGAGGGCCGAGGAAGAGATTTCCAACTATTA 1668
Qy 540 sAla1yThrPhe1eAlaLysG1uVal1SerThrThrG1uArGThrThr1eupheLysAspLeuG1 560
Db 1669 AGGTACTTCATAGCTTAAGAAAGTGTCTACCCAGGCAACATATCTGAAGATCTCGA 1728
Qy 560 uVal1eThrSerTrpPheG1nSerThrVal1SerLYsG1uAspAlaMetProG1uAla1e 580
Db 1729 AGTTATCACTCTGTGTTTCAAGCAGACAGTGAAGCAAGAGCGCATGCCGGAACACT 1788
Qy 580 uLYsSerLeu1ePheProAspPheG1uProLeuN1s1yLysPheN1s1Th1AspPheLeu 600
Db 1789 GAAGATCTCATATTTCCCAATTTTGAACCTTTGCAAAATTTCAATATTTTCTCAA 1848
Qy 600 sG1u1eG1uG1nArG1eUAla1eupThrG1uG1yArGAspAspAlaG1n1LeArAspTyr 620
Db 1849 GGAATTTAGCAAGCTTGCCCTGTGGAAAGCCGCTCAAAATGCCAAATTCAGATTA 1908
Qy 620 rG1nArG1eG1yAspVal1MetLeuLYsAspN1eG1nG1yMetLYsN1s1eUAla1aN1 640
Db 1909 CCAAGAAATCCGGAGATGATGCTGAAGAAANTTCAAGGCGATGAACACTGCGCGCTCA 1968
Qy 640 sLeuTh1yN1s1eSerG1uA1a1eug1uA1a1eug1uAspG1y1eLysSerSerArGAr 660
Db 1969 CTTGTGAAGACAGAGGAGGCTTGAGAGCCCTGAGAAATGGAATCAAGAGCTCCGGCGG 2028
Qy 660 g1eug1uAspPheCyArGAspPheG1uLeuG1nLYsVal1CyArTYrLeuProLeuAspN1h 680
Db 2029 GCTGGAGAACTTTGAGAGACTTGAAGCTGAAGAGGTGTGTACTTACCTGCTCAACAC 2088
Qy 680 rPheLeuLeuArGProLeuN1s1eArG1eMetN1s1yTrLYsG1nVal1eug1uArGLeuCy 700
Db 2089 CTTTCCTCTGCGGCGCACTGACCCGCTCATGCACTCAAGCAGCTCTGAGAGGCGCTGTG 2148
Qy 700 sLYsN1s1e1eProProSerN1s1eAlaAspPheArGAspCyArGAlaAla1eUAlaG1u1 720
Db 2149 CAACACACACCCGCGAGCAGCCGCACTTCAAGGAGCTCCGAGCGCTTTGGCAGAGAT 2208
Qy 720 eThrG1uMetVal1aG1nLeuN1s1eG1yThrMet1eLYsMetG1uAspPheG1nYsLe 740
Db 2209 CACGAGATGTGTGCAAGCTCCACGTTACGATGATCAAGATGGAATTTCCAGAAAGCT 2268

```

Oy 740 uHiegluLeuYslysaPleuilegYileaspaNleuValProglYarGluPh 760
Db 2289 GCAAGACTTCAGAAAGATTGATTGGCATTCGATCAATCTTGAGTCCGGAGAGGAGACT 2328
Oy 760 eilearGleuGlYserLeuSerlyLeuSerGlylyeGlYleuGlnGlnatGmetPhePh 780
Db 2239 CATCGCTGTGGGAGCCTCAGCAAGCTCTCGGGAGAGGGGCTCCAGCAGCGCATTTCTT 2388
Oy 780 eleuPheaaNaaPValleuLeuYrThSerarGlyleuThralaserAaenGlnPheLy 800
Db 2389 CCGTTTCAACACACGCTCGCTATACACAGCGGGGGCTGACGGGCTTCAATCAAGTTTAA 2448
Oy 800 aValHieglYleuPleuPleuYrGlymeThrllegluuGuserGluuPaGluTrgl 820
Db 2449 AGTCACGGGAGGCTCCGCTTATGGCATGACATTAAGAGAGAGAGAGAGAGAGAGAGAG 2508
Oy 820 YValProHiasGylleuThrlleuargGlyGlnargGlnserileleValalaserSe 840
Db 2509 GGTGCCCACTGGCTGACCTCCGGGGCCAGGGGCGATTCATCATCTGGCCGCAAGTTC 2568
Oy 840 rArSerGluuGluYrTrpValGluAapllleglmeAlalaeapleuAlaglyly 860
Db 2569 TCGGTCCAGATGGAGAGTGGGTTGAGAGACATCCAGATGGCCATTCAGCTGGCGAGAA 2628
Oy 860 sSerSerProHlaPProGluPheleuAlaserSerProPaPaPaPaPaPaPaPaPaPa 880
Db 2629 GAGACGAGCGCCCGCTGAGTTCTTGCCAGCGAGCGCCCTGACCAAGTCCCTCGA 2688
Oy 880 PglualathralalaePglngluSerGluuPaPaPaPaPaPaPaPaPaPaPaPaPaPaPa 900
Db 2689 TGAAGCCACCGCGGTGACAGAGTCAAGAGATACCTGAGCGCTGCCACATTCCT 2748
Oy 900 uGluarGlnalaeProHiasrGlyAaenThmeValHieValCyTrpPhisargAaenTh 920
Db 2749 GAGACGCGAGCGCCCGCGACCGCGCAACATGCTGACGCTGCTGCGACCGCGAACAC 2808
Oy 920 rSerValsermeValaPhePheSerilealValGluuAaenGluuSerGlyAaenLeu 940
Db 2809 CAGGCTCTCCATGGGAGCTTCAGATGACAGTGAAGATCAAGTGTGGAAGACTGCT 2868
Oy 940 uArGlyPheLyAaenSerAaenGlyTrpGlnlyleuTrpValPheThraaPheCy 960
Db 2869 GAGGAATTCAAAACAGCAACGGGTGCAAGCTGTGGGTGCTTCAAACTTTG 2928
Oy 960 eLeuPhePheYrlySerHieglNaPaPaNhiePProleuAlaserleuPleuLeuGl 980
Db 2929 CCGTTCTTCTAACAATCACACAGCAACATCATCCCTGCGAGCTGCTGCTCGG 2988
Oy 980 YTrSerleuThrllePProSerGluuSerGluuAaenleGlnlysaPlyrValPheLyLe 1000
Db 2989 CTACTCGCTCACATCCCTCTGAGTCCGAGAAACCTCAGAAAGACTACGTTCAAGCT 3048
Oy 1000 uHiePheLySerHieValYrTrPheArGAlagluSerGluYrThPheGluarGTr 1020
Db 3049 GCATTCAGAGTCCACGCTACTACTTAAGGGCCGAAAGCAAGTTCGAAAGGTG 3108
Oy 1020 pMeGluValleleargSerAlaThSerSerAlaserArpProHieValleuSerHieLy 1040
Db 3109 GATGGAAGTGAATCGCGAGTGCACACAGCTGCTCGGAGCCCACTGTGAGGACCA 3168
Oy 1040 eGluSerleuValYr 1045
Db 3169 AGAGTCTCTGTGTAT 3184

```

DE Human tumour-associated antigenic target (TAT) cDNA sequence #1621.
 XX human; tumour-associated antigenic target; TAT; cytostatic; gene therapy;
 KW cancer; cell proliferative disorder; gene; ss.
 XX Homo sapiens.
 OS WO2004060270-A2.
 XX
 PD 22-JUL-2004.
 XX
 PF 15-OCT-2003; 2003WO-US029126.
 XX
 PR 18-OCT-2002; 2002US-0418988P.
 XX
 PA (GETH) GENENTECH INC.
 PA (WO/ID/) WO T D.
 PA (ZHOU/) ZHOU Y.
 XX
 PI Wu TD, Zhou Y;
 XX
 DR WPI; 2004-534300/51.
 XX
 PT New nucleic acid molecule and encoded polypeptide, for diagnosing,
 PT preventing or treating cell proliferative disorders such as cancer.
 PS
 XX Claim 1; SEQ ID NO 1621; 5504pp; English.
 XX
 CC The present invention describes an isolated tumour-associated antigenic
 CC target (TAT) nucleic acid comprising: (a) any of 4622 nucleotide
 CC sequences (see SEQ ID NO:1 to 4622); (b) the full-length coding region of
 CC (a); (c) the complement of (a) or (b); (d) a sequence that has 80%
 CC sequence identity to (a)-(c); or (e) a sequence that hybridises to (a)-
 CC (c). Also described: (1) an expression vector comprising the above
 CC nucleic acid; (2) a host cell comprising the above expression vector; (3)
 CC a process for producing a polypeptide; (4) an isolated polypeptide
 CC comprising: (a) an amino acid sequence encoded by any of the above
 CC nucleotide sequences; (b) an amino acid sequence encoded by the full-
 CC length coding region of the above nucleotide sequences; or (c) a sequence
 CC having at least 80% identical to (a) or (b); (5) a chimeric polypeptide
 CC comprising the above polypeptide fused to a heterologous polypeptide; (6)
 CC an isolated antibody that binds to the above polypeptide; (7) a process
 CC for producing the antibody; (8) an isolated oligopeptide that binds to
 CC the above polypeptide; (9) a tumour-associated antigenic target (TAT)
 CC binding organic molecule that binds to the above polypeptide; (10) a
 CC composition of matter comprising the above (chimeric) polypeptide,
 CC antibody, oligopeptide or TAT binding organic molecule, in combination
 CC with a carrier; (11) an article of manufacture comprising a container and
 CC the composition of matter contained within the container; (12) methods of
 CC inhibiting the growth of a cell that expresses the above protein, where
 CC the growth of the cell is at least in part dependent upon a growth
 CC potentiating effect of the above protein; (13) a method of
 CC therapeutically treating a mammal having a cancerous tumour comprising
 CC cells that express the above protein; (14) a method of determining the
 CC presence of a protein in a sample suspected of containing the protein
 CC described above; (15) methods of diagnosing the presence of a tumour in a
 CC mammal; (16) a method for treating or preventing a cell proliferative
 CC disorder associated with increased expression or activity of the above
 CC protein; and (17) a method of binding an antibody, oligopeptide or
 CC organic molecule to a cell that expresses the protein described above.
 CC The TAT sequences have cytostatic activities, and can be used in gene
 CC therapy. The composition and methods are useful for diagnosing,
 CC preventing or treating cancer. The composition is also used for preparing
 CC a medicament for the therapeutic treatment or diagnostic detection of a
 CC cell proliferative disorder or cancer. The present sequence represents a
 CC human TAT cDNA sequence from the present invention.
 XX
 SQ Sequence 3997 BP; 998 A; 1132 C; 1068 G; 799 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 1,31e-189 Length: 3997
 Score: 2941.50 Matches: 580
 Percent Similarity: 70.0% Conservative: 162

Best Local Similarity: 54.7% Mismatches: 261
Query Match: 53.8% Indels: 57
DB: 12 Gaps: 14
US-09-555-342B-2 (1-1045) x AD084807 (1-3997)

Qy 1 MetGlyLeuIleGluGlnArg-----ProThrProGlySerArgLeuGlyAla 16
Db 118 ATGGGGAGATGAAAGAACTACAGAGTCCTCGACATCGAGGAGATCGCGCTGGGCGC 177
Qy 17 ProGluAsnSerGlyIleSerThrLeuGluArgGlyGlnIlyProProProThrProSer 36
Db 178 CAGACCCCTGGAGATTAGACCCCTTGAGCGCTGGGAGACTCTCTTCCCGAGATGCAA 237
Qy 37 GlyLeuValSerIleValIleGluMetLeuAspAspThrGlnIlyAlaPheGluVal 56
Db 238 GAGAGACCTGGACCTCAGAGTAAAGCTGTGGACAAACCATGGAAATATTTGACATT 297
Qy 57 ProGlnArgAlaProGlyIlyValIleuLeuAspAlaValCysAsnHisLeuAsnLeuVal 76
Db 298 GAGCCTAAATGCGATGGCCAGGTATTACTGACAAATGTGGGAAGCGTTTAACTGGTA 357
Qy 77 GluGlyAspTyrPheGlyLeuGluPheProAspHisIlyValIleThrValTyrLeuAsp 96
Db 358 GAATGTGACTACTTCGGGATGGAGTTTCAAAATACTGACTGATTTGGCTTGGAA 417
Qy 97 LeuLeuAspProIleValIlyValIleArgArgProIlyValIlyValIlyPheVal 116
Db 418 CCTATGAAACCATTTAGCCAAATACGAAAGCCAAAGATGGTGCTTCGCTAGCT 477
Qy 117 ValIlyPhePheProProAspHisThrGlnLeuGlnIlyGluLeuThrArgTyrLeuPhe 136
Db 478 GTAAATATTTTTCACCTGATCTCGTCAGCTACAAAGAAATATACAAAGTACTTGT 537
Qy 137 AlaLeuGlnValIlyGlnAspLeuAlaGlnIlyArgLeuThrCysAsnAspThrSerAla 156
Db 538 GCCTTGCATTAAGAGACCTGCTGGAAGAGGTTTGAACCTGTGTCGACACCAACGCG 597
Qy 157 AlaLeuLeuIleSerHisIleValGlnSerGlnIleGlyAspPheAspGlnAlaLeuAsp 176
Db 598 GCCCTTTCACGTCCTCCATCTCTGCAAGTCGAAATAGAGATTCAGTAAACCGCTGAC 657
Qy 177 ArgGlnHisLeuAlaIlyAsnIlyTyrIleProGlnIlyAspAlaLeuGlnIlyAspIle 196
Db 658 CGAGAGACCTCAAGAGGAAAGATATTGCTGCGCAGACGACGCTCTTGAGAAATATA 717
Qy 197 ValGluPheHisIleAsnHisIleGlyGlnThrProAlaGlnSerAspPheGlnLeu 216
Db 718 CTGAATTCATCAAGAGACGTCGGCCAGACCTGCTGATCGGATTTCCAGGTGCTC 777
Qy 217 GlnIleAlaArgArgLeuGluMetTyrGlyIleArgLeuHisProAlaIlyAspArgGlu 236
Db 778 GAATTCCTGAAAGATTGAAATGTACGGCATCGATTTCACATGCGCTTTCGACAGGAAA 837
Qy 237 GlyThrIlyHisLeuLeuAlaValAlaAsnThrGlyIleLeuValPheGlnIlyPheThr 256
Db 838 GAAACCAAGATTCACTGCGACGATTCCACATGGGTGTACTCGGTTCAGGGGACACACC 897
Qy 257 IlyGlnLeuAlaPheAsnThrAlaIlyValIlyArgIlyLeuSerPheIlyValGlyPhe 276
Db 898 AAATATCAACTTTCATCGTCCMAAGTCCGTAACTTAAGCTTCAAGAGAAAGATTTT 957
Qy 277 LeuIleValSerArgProAspAlaAsnSerAlaTyrGlnAspThrIleGluGluPheLeuMet 296
Db 958 CTATATCAACTTTCATCGAGAGTTTCATGACCTTACAGAGACATTTAGATTTTGTG 1017
Qy 297 AlaSerArgAspPheCysIlySerPheTyrIlyIleCysValGlnHisIleAlaPhePhe 316
Db 1018 GGTAGTGAATGATGATTAAGAACTTCTGMAAGATTGTGTGAGATATCACACTTTT 1077
Qy 317 ArgLeuPheGluGluProIlyProIlyProIlyProValLeuPheSerArgIlySerSer 336
Db 1078 AGACTTTTGAACCACTTAAGCCAAAGCAAAAGCCGCTTCTTCAACCGGGGCTCTCC 1137

Qy 337 PheArgPheSerGlyArgThrGlnIlyGlnValIleuAspTyrValIlyGlnIlyGlyHis 356
Db 1138 TTCAATACAGTGGAAAGAACTCGAAACAACTATGATTTATTTCAAGACATGGAAATG 1197
Qy 357 IlyLeuValGlnPheGluArgIlyHisIleValHis---SerIleArgSerLeu--- 374
Db 1198 AAGAGATTCATATGAAAGAGACAGACAGACCAACGTCCTGAGCTGTGACT 1257
Qy 375 AlaSerGlnProThrGluLeuAsnSerGlyValLeuGlnGlnSerGlnIlySer 394
Db 1258 GCAGACCTACCAAAAGAG-----AGC 1278
Qy 395 LeuThrPheGlyValIlyAlaGlnSerProGlyGlyGlnSerCysArgArg----- 411
Db 1279 ATCTCATTCGCCAGAGGATTTGAGAGACTCTGCTCCCATCTTCAGAGAAATGCTTTTAC 1338
Qy 412 ---GlyIlyGluProIlyValSerAlaGlyIlyProGly----- 423
Db 1339 TCGCTCTCCCTCCACTCTGTCCTGCTGCGCTGCGAGATTAAAGACAGACAGCAGC 1398
Qy 424 ---SerHisProSerProAlaProArgArgSerProAlaGlyAsnIlyGlnAlaAsp 441
Db 1399 TCCCTCACAGATCCGAGGTTTCTACGTCAAGAGTCCAGCTGCAAGAGCGCAGTGA 1458
Qy 442 GlyAlaAlaSerAlaProThrArgIlyGlnIlyGlnValIlyAspArgThrGln 461
Db 1459 GCAGTGCTGAGAGCCCGAC-----AACCATCG 1488
Qy 462 SerIlyProGlnProPro-----GlnProSerThrGlySerLeuThrGlySerPro 478
Db 1489 GCCAGACCCCTCGAGGCCCCCGCATCCAGCTGTCAGGCTTTCACGAAGAGTCCT 1548
Qy 479 HisLeuSerGlnLeuSerValAsnSerGlnIlyGlyValAlaProAlaAsnValThrLeu 498
Db 1549 CAGCTTCTCCTCCAGCGGAAAGACCCCTGAGTGTGAGCTGCA-----TTT 1599
Qy 499 SerProAsnLeuSerProAspThrIlyGlnAlaSerProLeuIleSerProLeuAsn 518
Db 1600 CAGTCCCTTTGGGCCCACTGAAACAGGGCTCATCCCATCTCTGAGCCCTGCTCACT 1659
Qy 519 AspGlnAlaCysProArgThrAspArgIlyAspGlnIlyArgArgIlyAspPheProThr 538
Db 1660 GATGCTGGCGGAGCCGGAATGACTGCGAG---GAACCCAGACAAAGCCGCTGCA 1716
Qy 539 AspIlyAlaTyrPheIleAlaIlyGlnValSerThrThrGlnArgThrTyrLeuIlyAsp 558
Db 1717 GACGAGGCTTACTTCAATAGTCAAGAGATTCTCGCTACAGAAACATACCTCAAGAT 1776
Qy 559 LeuGlnValIleThrSerTyrPheGlnSerThrValSerIlyGlnAspAlaMetProGlu 578
Db 1777 TTGAAGATTATTCGTCGTGTTCCGACGCGAGTGTGAAGAGAGCGCATGCGCTGCG 1836
Qy 579 AlaLeuIlySerLeuIlePheProAsnPheGluProLeuHisIlyAspPheHisThrAsnPhe 598
Db 1837 ACTCTATAGCGCTGCTCTTTCACATGATCCCATGATATGATGATTCACAGAGCTTC 1896
Qy 599 LeuIlyGlnIleGlnIlyArgLeuAlaLeuThrProGlnIlyArgSerAsnAlaGlnIleArg 618
Db 1897 CTGCGGAGGTGAGAGAGGCTGCGACTCTGGAAGGGGCGCTCCAAAGCCACACAAAA 1956
Qy 619 ---AspTyrGlnArgIleGlyAspValMetLeuIlyAsnIleGlnIlyMetIlyHisLeu 637
Db 1957 GCGAGTCATCAAGATCCGGGACATCTGCTCAGAAACATGCGCAGATTAAAGAGTTT 2016
Qy 638 AlaAlaHisLeuThrIlyHisSerGlnAlaLeuGlnAlaLeuGlnAsnIlyIlySer 657
Db 2017 ACCAGCTACTTCCAAAGACATGACGAGTCTTAAAGAACTGGAAGAAAGCTACCAAGC 2076
Qy 658 SerArgArgLeuGlnAsnPheCysArgAspPheGlnLeuGlnIlyValCysTyrIleuPro 677
Db 2077 TGTAAAGATTGAGGCACTGTACAAAGAGTTTGAAGCTGCAAGAGGTCTGCTACTGCT 2136

```

Qy 678 LeuAnThrPheLeuLeuArgProLeuNH1aArgLeuMeCH1eTYrLYeGlnValLeuGlu 697
Db 2137 CTCAACAGCTTCCTGCTGAAGCCATCCAGCGGCTGCTGCACTACCGCTGCTGGCC 2136
Qy 698 ArgLeuCyVlyNH1aNH1aProProSerNH1aLaerPheArgSerCyAArgAlaAlaLeu 717
Db 2197 CGCCTATGCGGACATTACAGCCCGGCGACCATGACTACGCTGACTGCGCATGACGCCCTG 2256
Qy 718 AlaGluIleThrGluMetValAlaGlnLeuNH1aGlyThrMetC11eLYeMetGluAsnPhe 737
Db 2257 AAAGCCATCACAAGAGTACCAACCACTACACGACATTCATCCGCTGGAGAACCTG 2316
Qy 738 GlnLYeLeuNH1aGlnLeuLYeLeuPheLeuIleGly11aAspAlaValProGly 757
Db 2317 CAGAACCTAAAGGAGCTGACCGGGACCTGGTGGGCAAGAACCTATGCTCTGGC 2376
Qy 758 ArgGluPheIleArgLeuGlySerLeuSerLYeLeuSerGlyLYeGlyLeuGlnArg 777
Db 2377 AGGAGATTCACTCGTAGGGCTGCTTCAAAAGCTCAACAAAGAGGCTGACAGAGG 2436
Qy 778 MetPheLeuPheAsnArgValLeuLeuTYrThrSerArgLYeLeuThrAlaSerAsn 797
Db 2437 ATGTTTTTCTGTTCTCAGATATGTTGCTGACACAAAGAGATTGACAGGACGACG 2496
Qy 798 GlnPheLYeValNH1aGlyGlnLeuProLeuTYrGlyMetThrIleGlnGluSerGluAsp 817
Db 2497 CACTTCGGGATCCGGGCTCTTCCCTCCAGGCAATGCTGCGTGAAGAAAGATGATAC 2556
Qy 818 GlnTYrGlyValProNH1aCyVleuNH1aLeuArgLYeGlnArgGlnSerIle11aValAla 837
Db 2557 GAGTGGTCTGTTCCACACTGTTTCAACCATTCACGCGGCTGAGAAACAAATGCTGGGCA 2616
Qy 838 AlaSerSerArgSerGlyMetGlnLYeTYrValGluAspIleGlnMetAla11aAspLeu 857
Db 2617 GCCGACCTCGGCTGGAGAAAGAGAGTGTGACCTGCAACTCCGCTATCCAAAGCA 2676
Qy 858 AlaGluLYeSerSerSerProAlaProGluPhe-----LeuAlaSerSerPro 873
Db 2677 GCCAAGAGTGGCGGTGACAGCGCCCTGCTGACTGCCAGGCGGCACGTGTGACTGCTGCC 2736
Qy 874 ProPheArgLYeSerProArgGlyAlaThrAlaAlaArgGlnLYeSerGluAspArgLeu 893
Db 2737 CCC-----AGATCCCCCAAGAG---GTATCTCTGAGAGCAGAGTCAAGAAATGATGCT 2787
Qy 894 SerAlaSerArgThrSerLeuGlnArgGlnAlaProNH1aArgLYeAsnThrMetValNH1a 913
Db 2788 CGGGGTCTCGGACCTCTGAGAGGGGATGGCCAGCACCGGGCCAACACCAATGCAC 2847
Qy 914 ValCYeTYrNH1aArgAsnThrSerValSerMetValAspPheSerIleAlaValGluAsn 933
Db 2848 GTGTGCTGTGTAACCGAAACACAGGCTGTCCAGGGCAACACCAATGCACTGTGAAAGAC 2907
Qy 934 GlnLeuSerGlyAsnLeuLeuArgLYePheLYeAsnSerArgGlyTYrGlnLYeLeuTYr 953
Db 2908 CAGCTTCAAGATATCTGCTAAAGAAAGTTCAAAACAGTATGCTGGCAGAAAGCTCTGG 2967
Qy 954 ValValaPheThrAsnArgPheCyVleuPhePheTYrLYeSerNH1aGlnAspAlaPheLeu 973
Db 2968 GTGCTTTTACCAATCTGTGTTGTTCTTCAAAATCATCAAGATGATACCACTG 3027
Qy 974 AlaSerLeuProLeuLeuGlyTYrSerLeuThrIleProSerGluSerGluAsnIleGln 993
Db 3028 GCCAGCTTCGCGCTGGGTGACAGCTGACATCCCAAGAGGCGGATGSGATACAC 3087
Qy 994 LYeArgTYrValaPheLYeLeuNH1aPheLYeSerNH1aValTYrTYrPheArgAlaGluSer 1013
Db 3088 AAAGACATATGTTTCAAGCTCAGTCAAAATCCCAAGCTTACTTCCGGGCTGAGAGC 3147
Qy 1014 GlnTYrThrPheGlnArgLYeMetGluVal11aArgSerAlaThrSerSerAlaSerArg 1033
Db 3148 AAGTACATTTTGAAGGTGAGTATGATCCAGGGGCGCAACAGCTCAGCGGAGG 3207

```

```

ADRs5675
ID ADRs5675 standard; DNA; 3997 BP.
XX
AC ADRs5675;
XX
DT 21-OCT-2004 (first entry)
XX
DE Breast cancer prognosis marker #1536.
XX
KM ds; breast cancer; prognosis; gene expression; diagnosis.
XX
OS Homo sapiens.
XX
PN M02004065545-A2.
XX
PD 05-AUG-2004.
XX
PF 15-JAN-2004; 2004WO-US001100.
XX
PR 15-JAN-2003; 2003US-00342887.
XX
PA (ROSE-) ROSETTA INPHARMATICS LLC.
XX
PE (MECA-) NETHERLANDS CANCER INST.
XX
PI Van't Veer LJ, He Y;
XX
DR WPI; 2004-593473/57.
XX
PT Classifying a breast cancer patient according to prognosis comprises
PT determining the similarity between the level of expression of each of
PT five genes in a cell sample taken from patient, to control levels.
XX
PS Disclosure; SEQ ID NO 1536; 226bp; English.
XX
CC The invention relates to a method of classifying a breast cancer patient
CC according to prognosis by determining the similarity between the level of
CC expression of each of five genes for which markers are listed in the
CC specification, in a cell sample taken from the breast cancer patient, to
CC control levels of expression for each respective five genes to obtain a
CC patient similarity value. The methods are useful for classifying a breast
CC cancer patient according to prognosis. Kits and computer program products
CC are useful for data analysis using the diagnostic, prognostic and
CC statistical methods of the invention. This sequence corresponds to a
CC marker used in the method of the invention.
XX
SQ Sequence 3997 BP; 998 A; 1132 C; 1068 G; 799 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 1,31e-189 Length: 3997
Score: 2941.50 Matches: 580
Percent Similarity: 70.0% Conservative: 162
Best Local Similarity: 54.7% Mismatch: 261
Query Match: 13.8% Indels: 57
DB: Gaps: 14
US-09-555-342B-2 (1-1045) x ADRs5675 (1-3997)
Qy 1 MetGlyGlnIleGlnArg-----ProThrProGlySerArgLeuGlyAla 16
Db 118 ATGGGGGAGATAGAGAAAGACATACAGAGTCTGCACTGACGGGATGCGCTGGGGTGGCC 177
Qy 17 ProGluAsnSerGlyIleSerThrLeuGluArgGlyGlnLYeProPheProThrProSer 36
Db 178 CAGAACCTTGGAGATTGACACCTTGAGACCTGGGCAAGCTCTTGGCCAGAAATGCAA 237
Qy 37 GLYeLeuValSerIleLYeGlnMetLeuAspArgThrGlnGlnAlaPheGluVal 56
Db 238 GAGAAAGCACTGCACTCAGAGTAAAGCTGTGCAACAAACCAATGAAATATTGACATT 297
Qy 57 ProGlnArgAlaProGlyLYeValLeuLeuAspAlaValaCyAsnNH1aLeuAsnLeuVal 76
Db 298 GAGCTAAATGCGATGGCCAGATATTACTGACACAAGTGTGAAAGCGTTTAAACCTGGTA 357

```


Dh 2437 ATGTTTTTCTGTTCTGATATGTTGCTGTACACAAAGAGATTGACGAGCACC 2436
Qy 798 GlnPheIyValIhAGIyGlnLeuProLeuTyrgIyMetThrIleGIuGluSerGIuAsp 817
Dh 2497 CACTTCGGGATCCGGGGCTCTCTCCCTCCCAAGGACGTGGTGAAGAAAGTATAC 2556
Qy 818 GIuTrpGIyValIProhIaCyLeuThrLeuArgGIuGlnArgIInserIleIleValIa 837
Dh 2557 GAGTGGTGTGTTCCACACTGTTTCCATTCACCGGGCTCGACAAAACATGTGTGCA 2616
Qy 838 AlaSerSerArgSerGIuMetGIuIyTrpValGIuAspIleGIuMetAlaIleAspLeu 857
Dh 2617 GCCAGCACTCGGTGAGAAAGAAAGATGTGTGACCTGAACCTCCGGCATCCAAACA 2676
Qy 858 AlaGIuIySerSerSerProAlaProGIuPhe-----LeuAlaSerSerPro 873
Dh 2677 GCCAAGAGTGGCGGTGACACCGCCCTGCACTGCGACGCCCACTGTGTGACCTGTC 2736
Qy 874 ProAspAnIySerSerProAspGIuAlaThrAlaIleAspGIuIySerGIuAspAspLeu 893
Dh 2737 CCC-----AGATCCCCCAAGCAG--GTATCTGTGGAGCAGAGTCCAGAAAGATGATCT 2787
Qy 894 SerIaSerArgThrSerLeuGIuArgIuAlaProhIaArgGIuAspThrMetValhIs 913
Dh 2788 CGGGGTGTCGACAGTCCCTGAGAGGGCATGGCCAGCACCGGGCCAAACACCAATGCAC 2847
Qy 914 ValIyTrpIhIaAGIyValhIaSerValSerMetValaAspPheSerIleAlaValGIuAsn 933
Dh 2848 GTGTGCTGTGTCACCGAACAACACAGCGTGTCCAGGGCAGACCAAGTGCAGCTGTGAGAAC 2907
Qy 934 GlnIySerGIyAsnLeuLeuAlaGIyPheIyAsnSerAsnGIyTrpGIuIyLeuTrp 953
Dh 2908 CAGCTTTCAGGATATCTGCTTAAGAAAGTTCAAAACATGCTAGGCTGCAAGAGCTTGG 2967
Qy 954 ValValPheThrAspPheCyLeuPhePheTyrlYySerhIsgIuAspAnhIaProLeu 973
Dh 2968 GTGCTCTTACCAATCTCTGTTTGTCTTCAAAATCACTACGAGTACCACTG 3027
Qy 974 AlaSerLeuProLeuGIyTrpSerLeuThrIleProSerGIuSerGIuAsnIleGIu 993
Dh 3028 GCCAGCTCCCGCTGGCTGGGTACAGCCTGAGCATCCCGAGGAGCGCATGACATAC 3087
Qy 994 LysAspTrpValIhPheIyLeuIhIaPheIySerhIaValIyTrpPheArgAlaGIuSer 1013
Dh 3088 AAAGACTATGTTCACAGCTCAAGTTCAAATCCAGCTTACTTCTCCGGGCTGAGAGC 3147
Qy 1014 GIuTrpThrPheGIuArgTrpMetGIuValIleArgSerAlaThrSerSerAlaSerArg 1033
Dh 3148 AGTATACATTTGAAGGTGATGAGGTGATCCAGGGGGCCAGAGCTCAGCCGGAGG 3207

RESULT 8
ADS64452 standard; DNA; 3997 BP.
AC ADS64452;
XX 16-DEC-2004 (first entry)
XX Human KIAA0793 gene.
XX Human; Interferon; IFN; gene; ds; KIAA0793.
XX Homo sapiens.
XX US2004185489-A1.
XX 23-SEP-2004.
XX 17-MAR-2004; 2004US-00802432.
XX 21-MAR-2003; 2003EP-00006263.
XX (CERT/) CERTA U.

PA (FOSE/) FOSER S.
PA (WEYE/) WEYER K.
PI Certa U, Foser S, Weyer K;
XX MPI; 2004-689187/67.
DR GENBANK; AB018336.
PT Determination of biological activity of compound which can modulate gene
PT transcription, involves contacting host with compound, determining
PT general transcriptional gene response of host, and quantitating gene
PT response induced by compound.
XX
PS Claim 8; SEQ ID NO 16; 53pp; English.
XX
CC The invention relates to a method for determining the biological activity
CC of a compound which can modulate gene transcription. The method involves
CC contacting a host with a compound, determining the general
CC transcriptional gene response of the host and quantitating the gene
CC response induced by the compound. This gene transcription assay method is
CC useful for the determination of the biological activity of a compound
CC which can modulate gene transcription. The present sequence is the human
CC KIAA0793 gene. This sequence is used to estimate the influence of the
CC pegylation site on the transcriptional activity of interferon (IFN) by
CC measuring this IFN-induced gene expression pattern of the melanoma cell
CC line ME15 via the oligonucleotide array technology.
XX
SQ Sequence 3997 BP; 998 A; 1132 C; 1068 G; 799 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 1,31e-189 Length: 3997
Score: 2941.50 Matches: 580
Percent Similarity: 70.0% Conservative: 162
Best Local Similarity: 54.7% Mismatches: 261
Query Match: 53.8% Indels: 57
DB: Gaps: 14
US-09-555-342B-2 (1-1045) x ADS64452 (1-3997)
Qy 1 MetGIyGIuIleGIuGIuArg-----ProThrProGIySerArgLeuGIyAla 16
Dh 118 ATGGGGGAGATAGAAAGAAACATACAGATCTCGACAGCTGACGGAGATGCGCTGTGGTCC 177
Qy 17 ProGIuAsnSerGIyIleSerThrLeuGIuArgGIyGlnIySerProProThrProSer 36
Dh 178 CAGACCCCTGTGGAGGTAGACCCCTGAGCCTGGGAGACTCTTGGCCCAAGATGCA 237
Qy 37 GIyIySerValIleSerIleIyLeuIleGIuMetLeuAspAspThrGIuAlaIhPheGIuVal 56
Dh 238 GAGAAAGCACTGCACCTCAGAGTAAAGCTGTGCAACAACCAAGAAATTTTGACATT 297
Qy 57 ProGIuArgAlaProGIyIyValIleuAspAlaValIyAspAnhIaLeuAsnLeuVal 76
Dh 298 GAGCTTAAGTCGATGCGACAGGTATTTACTGACACAAAGTGGAAAGCGTTTAAACCTGGTA 357
Qy 77 GIuGIyAspTrpPheGIyLeuGIuPheProAspIhIyIySerIleThrValITrpLeuAsp 96
Dh 358 GAATGTACTACTTCGGAGATGAGAGTTTCAAAATCTAGCTTCACTGAGATTGGCTTGAA 417
Qy 97 LeuLeuIySerProIleValIyGIuIleArgArgProIyIhIaValIyValIyPheVal 116
Dh 418 CCTATGAACCCATCATTAGCAAAATACGAAGGCCAAAGATGTGTGCTTGGCTGACT 477
Qy 117 ValIyAspPhePheProProAspIhIaThrGIuGIuGIuGIuLeuThrArgTrpIyLeuPhe 136
Dh 478 GTAAATTTTTCACCTGATCTCTGTGACGTACAAAGAAATATACAAAGATCTTGT 537
Qy 137 AlaLeuGIuValIyGIuAsnIleuAlaGIuGIyArgLeuThrCyAspAspThrSerAla 156
Dh 538 GCTTGCACTTAAAGAGACCTGCTGAAAGCGTGTGACCTGTGACACCAACGAGCG 597
Qy 157 AlaLeuLeuIleSerhIaIleValGIuIySerGIuIleGIyAspPheAspGIuAlaLeuAsp 176

Db 598 GCCCTTCTCAGTCCCATCTCTCTGACGTGGAAATAGAGATTACGATGAACGCTGGAC 657
 Qy 177 ArgGlnHisLeuAlaIysAsnLysTyrIleProGlnGlnAspAlaLeuGluAspLysIle 196
 Db 658 CGAGAGACCTCAAGTAGAAGATATTGCTGGCCAGACGACCTGCTTGAGAGATA 717
 Qy 197 ValGluPheHisIleAsnHisIleGlyGlnThrProAlaGluSerAspPheGlnLeu 216
 Db 718 CTGAATTCATCAGAGACAGCGTGGCCAGACACCTGCTGATGCGATTCCAGGTCTC 777
 Qy 217 GluIleAlaArgLeuGluMetTyrGlyIleArgLeuHisProAlaLysAspArgGlu 236
 Db 778 GAATTCCTCGAAAGTTGGAATGTACGGCATCGATTCCATCATGCTTCTGACAGGAA 837
 Qy 237 GATThrTyrIleAsnLeuAlaValAlaAsnThrGlyIleLeuValPheGlnGlyPheThr 256
 Db 838 GGAACCAAGATTCAGCTGGCAGTTTCCACATGGGATTACTGTTTCCAGGGACACAC 897
 Qy 257 LysIleAsnAlaPheAsnThrAlaLysValArgLysLeuSerPheLysArgLysArgPhe 276
 Db 898 AAATCAACACTTTCAGCTGTCGCAAGTCCGTAATTAAGCTTCAAGAGAAAGATT 957
 Qy 277 LeuIleLysLeuArgProAspAlaAsnSerAlaTyrGlnAspThrLeuGluPheLeuMet 296
 Db 958 CTATCAAACTTCATCCAGAGGTTTCATGACCTTACCGAGACACATTAGAAATTTGTTG 1017
 Qy 297 AlaSerArgAspPheCysLysSerPheThrLysIleCysValGlnHisIleAlaPhePhe 316
 Db 1018 GGTAGTAGAATGATATATGAACCTTGAGAGATTGTGTGGAGATACACCTTTT 1077
 Qy 317 ArgLeuPheGluGluProLysProLysProLysProValLeuPheSerArgLysSer 336
 Db 1078 AGACTTTGGACCACTTAAGCCAAAGAAAGCCGCTTCTTCAAGCGGGGCTCTCC 1137
 Qy 337 PheArgPheSerGlyArgThrGlnLysGlnValLeuAspTyrValLysGluGlyHis 356
 Db 1138 TTCAAGTACAGTGGAGAACTTCAGAAACAATAGTACATTATTCAAGACAGTGATG 1197
 Qy 357 LysLysValGlnPheGluArgLysHisSerLysIleHis--SerIleAspSerLeu-- 374
 Db 1198 AAGAGATTCCATATGAAGAAGGACAGACAGAACCCACACGTCCTTCAGCTTGACT 1257
 Qy 375 AlaSerGlnProThrGluLeuAsnSerGluValLeuGluGlnSerGlnSerThrSer 394
 Db 1258 GCAGACTTACCAAAACAG-----AGC 1278
 Qy 395 LeuThrPheGlyGlyAlaGluSerProGlyGlyGlnSerCysArgArg----- 411
 Db 1279 ATCTCATTTCCCGAGGATTTGAGGACCTGCTCCCATCTTCAGGAAATGCTTTTAC 1338
 Qy 412 ----GlyLysGluProLysValSerAlaGlyGluProGly----- 423
 Db 1339 TCGCTCTCTCCCTCCACTGCTGTCCTCTGCTGCTGACAGATTTAAGACAGACAGAC 1398
 Qy 424 ----SerHisProSerProAlaProArgArgSerProAlaGlyAsnLysGlnAlaAsp 441
 Db 1399 TCCCTACAGATCCCGAGTTTCTCTACGTCAGAGTCACGTCGACAGAGAGCCGAGGA 1458
 Qy 442 GlyAlaAlaSerAlaProThrGluGluGluGluValValLysAspArgThrGln 461
 Db 1459 GCAGTGGCTGAGAGCCCGAC-----ACACCATCG 1488
 Qy 462 SerLysProGlnPro-----GlnProSerThrGlySerLeuThrGlySerPro 478
 Db 1489 GCCCAGCCCTCGGGCCCCCGCAGCTCAGCTGTCAGGCTTCCACAGAAAGTCT 1548
 Qy 479 HisLeuSerGluLeuSerValAsnSerGlnGlyValAlaProAlaAsnValThrLeu 498
 Db 1549 CAGCTTCTCCCTCCACCGGAGAGAGCCCTGAGTGTGAGCCCTGCA-----TTT 1599
 Qy 499 SerProAsnLeuSerProAspThrLysGlnAlaSerProLeuLysSerProLeuAsn 518
 Db 1600 CAGGTGCTTTGGGCCAGCTGAACAGGCTCATCTCCCATCTCGAAGCCCTGTCTCAGT 1659

Qy 519 AspGlnAlaCysProArgThrAspArgLysAspGluGlyArgArgLysArgPheProThr 538
 Db 1660 GATGCTGGGGAGACCCGGAGTGACTGCGAG---GAGCCAGACACAAAGCCGCTTGC 1716
 Qy 539 AspLysAlaTyrPheIleAlaLysGluValSerThrThrGluArgThrTyrLeuLysAsp 558
 Db 1717 GACGAGCCCTACTTCACTAGTCAAAAGAGATTCTGCTTACAGAAAGAACTATCCCAAGAT 1776
 Qy 559 LeuGluValIlePheSerThrPheGlnSerThrValSerLysGluAspAlaMetProGlu 578
 Db 1777 TTAGAAATTATTAACCTGTGTGTTCCGACGCGAGTGTGAAGGAGACGCTAGCTGGC 1836
 Qy 579 AlaLeuSerSerLeuIlePheProAsnPheGluProLeuHisLysAspHisThrAsnPhe 598
 Db 1837 ACTGTATGACGCTGCTCTTCTCCAAATCATGATCCCATGTATGATTCACAGAGCTTC 1896
 Qy 599 LeuLysGluIleGluGlnArgLeuAlaLeuThrGluGlyArgSerAsnAlaGlnIleArg 618
 Db 1897 CTGCGGAGGTGGAGAGAGCTGCGCATCTGGGAAGGCCCCCAAGCCACACAAA 1956
 Qy 619 ---AspTyrGlnArgIleGlyAspValMetLeuLysAsnIleGlnGlyMetLysIleLeu 637
 Db 1957 GGCAGTCAATCAAGAAATCGAGGACATCTGCTCAGAAACATGCGCAATTAAAGAGTTT 2016
 Qy 638 AlaAlaHisLeuThrLysHisSerGluAlaLeuGluAlaLeuGluAsnGlyIleLysSer 657
 Db 2017 ACCAGCTACTCCAAAGACATGACGAGTCTTACAGAACTGGAAAGCTACCAAAAGC 2076
 Qy 658 SerArgArgLeuGluAsnPheCysArgAspPheGluLeuGlnLysValCysTyrLeuPro 677
 Db 2077 TGTAAAGATGGAGCAGTGTACAAAGATTGAGCTCAGAAAGCTGCTCTTGGCT 2136
 Qy 678 LeuAsnThrPheLeuLeuArgProLeuHisLysLeuMetHisTyrTyrGlnValLeuGlu 697
 Db 2137 CTAAACGTTCTGCTGTGAAGGCCATCCAGGCGCTCTGACTTACCGCTGCTGCGC 2196
 Qy 698 ArgLeuCysLysHisIleAspProSerHisIleAspPheArgAspCysArgAlaAlaLeu 717
 Db 2197 CGCTATGGGACATTAACAGCCCGGCGACATGATAGAGCTGACGTCAGCGCCCTG 2256
 Qy 718 AlaGlnIleThrGlyMetValAlaGlnLeuHisIleGlyThrMetIleLysMetGluAsnPhe 737
 Db 2257 AAGCCATCACAGAGGTGACCAACCACTACAGACATTTCTATCCGGGGAACCTG 2316
 Qy 738 GlnLysLeuHisGluLeuLysLysAspLeuIleGlyIleAspAsnLeuValProGly 757
 Db 2317 CAGAGCTAAAGAGCTGACGCGGACCTGTGTGGCATGAAGACTCATGCTCTGCGC 2376
 Qy 758 ArgGluPheIleArgLeuGlySerLeuSerLysLysSerGlyLysGlyLeuGlnArg 777
 Db 2377 AGGAGTTCATCCGTGAGGGCTGCTTCAAGCTCACCAAGAAAGGCGCTGACAGAGG 2436
 Qy 778 MetPhePheLeuPheAsnAspValLeuLeuTyrThrSerArgGlyLeuThrAlaSerAsn 797
 Db 2437 ATGTTTTTCTGTTCACATATGTGTTGCTGTACACAAAGAGAGTTCAGGAGGACG 2496
 Qy 798 GlnPheLysValHisGlyGlnLeuProLeuTyrGlyMetThrIleGluGluSerGluAsp 817
 Db 2497 CACTTCGGAATCGGGGCTCTTCTCCCTTCAAGGACATGCTGTGAAGAAATGATATAC 2556
 Qy 818 GluThrGlyValProHisCysLeuThrLeuArgGlyLysArgIleSerIleIleValAla 837
 Db 2557 GAGTGTCTGTTCACATGTTTCAACATCTACGCGGCTCAGAAACAAATCGTGTGCA 2616
 Qy 838 AlaSerSerArgSerGluMetGlyLysThrValGluAspIleGlnMetAlaIleAspLeu 857
 Db 2617 GCCAGCATCTCGGTGAGAAAGAGTGAATGTAAGTCTGAACCTCCGGATTCAGGCA 2676
 Qy 858 AlaGluLysSerSerSerProAlaProGluPhe-----LeuAlaSerSerPro 873
 Db 2677 GCCAAGTGGGGGTGACAGGCGCTTGCATCCAGCGCGACCTGTGTGCACTGCTCC 2736

QY 874 ProkaphenylSerProaspGluAlaThrAlaAlaaspGlnIuserGluaspLeu 893
 DB 2737 CCC-----AGATCCCCCAAGAG---GTATCTCGAGCGAGGTCAAGAGATGATGCT 2787
 QY 894 SerAlaSerArgThrSerLeuGluArgGlnAlaProHisArgGluAsnThrMetValHis 913
 DB 2788 CGGGGTGTCGACACTCCCTGAGAGGGGATGCGACAGCCGGGCAACACCAATGAC 2847
 QY 914 ValCysThrPheAlaAspAsnThrSerValSerMetValAspPheSerIleAlaValGluAsn 933
 DB 2848 GTGTGCTGGAGACCGGAACACAGAGGTGTCCAGGCAAGCCACATGACAGCTGTGAGAAC 2907
 QY 934 GlnLeuSerGluAsnLeuLeuArgLysPheLysAsnSerAsnGlyTTPGlnLysLeuTyr 953
 DB 2908 CAGCTTCAGAGATATCTGCTTAAAGAACTTCAAAACAGTCAATGCTGCGACAGACTCTGG 2967
 QY 954 ValValPheThrAsnPheCysLeuPhePheTyrLysSerHisGlnAspAsnHisProLeu 973
 DB 2968 GTGCTTTACCAACTCTGTGTTGTTCTTACAAAATCATATGAGATGACTACCACTG 3027
 QY 974 AlAserLeuProLeuLeuGlyTyrSerLeuThrIleProSerGluSerGluAsnIleGln 993
 DB 3028 GCCAGCTCCCGCTGGCTGGCTACAGCGTACATCCCGAGGAGCGCATGCGATGCATCAC 3087
 QY 994 LysAspTyrValPheLysLeuHisPheLysSerHisValTyrTyrPheArgAlaGluSer 1013
 DB 3088 AAAGACTATGTTTCAAGCTCCAGTCAATCCAGCTTACTTCTCCGGGCTGAGAGC 3147
 QY 1014 GluTyrThrPheGluArgTyrMetGluValIleArgSerAlaThrSerSerAlaSerArg 1033
 DB 3148 AAGTACACATTTGAAAGATGATGAGATGATGATGAGGAGGAGGAGGAGGAGGAGG 3207

RESULT 9

AAC98992
 ID AAC98992 standard; cDNA; 3094 BP.

XX AAC98992;

XX 09-MAR-2001 (first entry)

DE Human pancreatic cancer antigen nucleotide sequence SEQ ID NO:220.

XX Human: pancreas; pancreatic cancer; pancreatic cancer antigen; detection;
 KM diagnosis; identification; cytotoxic; neuroprotective; nootropic;
 KM immunomodulatory; relaxant; contraceptive; gynaecological;
 KM antiinflammatory; cardiant; gene therapy; chromosome mapping;
 KM linkage analysis; tissue identification; tissue typing; forensic; neutral;
 KM immune system; muscular; reproductive; gastrointestinal; pulmonary;
 KM cardiovascular; renal; proliferative; ss.

XX Homo sapiens.

XX MO200055320-A1.

PD 21-SEP-2000.

PF 08-MAR-2000; 2000MO-US005989.

XX 12-MAR-1999; 99US-0124270P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Ruben SM;

XX WPI; 2000-579444/54.

XX P-PSDB; AAB54227.

PT New nucleic acid that is a pancreatic cancer antigen for preventing,
 PT treating, or ameliorating a medical condition, particular pancreatic
 PT cancer, or for use in assays for diagnosing a pathological condition.

PS Claim 1; Page 664-665; 1379pp; English.

CC AAC98773 to AAC99231 encode the human pancreatic cancer associated
 CC proteins, called pancreatic cancer antigens, given in AAB54008 to
 CC AAB54466. The human pancreatic cancer antigens have cytostatic,
 CC neuroprotective, nootropic, immunomodulatory, relaxant, contraceptive,
 CC gynaecological, cardiant and antiinflammatory activities, and can be used
 CC in gene therapy. The polynucleotide and proteins can be used for
 CC preventing, treating, or ameliorating a medical condition or in assays
 CC for diagnosing a pathological condition or a susceptibility to one in a
 CC subject. Binding partners to the proteins and the activity of the
 CC proteins can be identified. The pancreatic cancer antigens can be used to
 CC detect, treat or prevent pancreatic disorders, especially cancer.
 CC Agonists and antagonists to the antigens can be screened for. The
 CC pancreatic cancer antigen polynucleotides can be used to design nucleic
 CC acid hybridisation probes that can be used in chromosome mapping, linkage
 CC analysis, tissue identification and/or typing and a variety of forensic
 CC and diagnostic methods. The proteins can be used to generate antibodies
 CC which are used to purify, detect and target the polypeptides, including
 CC both in vivo and in vitro diagnostic and therapeutic methods. The
 CC proteins can be used to treat or prevent neural, immune system, muscular,
 CC reproductive, gastrointestinal, pulmonary, cardiovascular, renal or
 CC proliferative disorders. AAC99232 to AAC99240 and AAB54467 represent
 CC sequences used in the exemplification of the present invention

XX Sequence 3094 BP; 736 A; 849 C; 755 G; 749 T; 0 U; 5 Other;

Alignment Scores:

Score: 1.22e-159 Length: 3094
 Percent Similarity: 2498.00 Matches: 479
 Best Local Similarity: 99.6% Conservative: 0
 Query Match: 45.7% Mismatches: 2
 DB: 3 Indels: 1
 Gaps: 0

US-09-555-342B-2 (1-1045) x AAC98992 (1-3094)

QY 565 TrpPheGlnSerThrValSerLysGluAspAlaMetProGlnAlaLeuLysSerLeuIle 584
 DB 63 TGGTTTCAGAGCAGATGAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 122
 QY 585 PheProAsnPheGluProLeuHisLysPheHisThrAsnPheLeuLysGluIleGln 604
 DB 123 TTCGCAATTTTGAACCTTGGACAAATTCATACATTAATTTTCAAGGAATTTGAGCAA 182
 QY 605 ArgLeuAlaLeuTrpGlnGlyArgSerAsnAlaGlnIleArgAspTyrGlnArgIleGly 624
 DB 183 CGACTTGCCCTGGAG 242
 QY 625 AspValMetLeuLysAsnIleGlnGlyMetLysHisIleLeuAlaHisIleLeuTrpLysHis 644
 DB 243 GATGTCAATGCTGAAGAACTTCAAGGCGATGAGACCTGCGGCTCACTGTGAGAGCAC 302
 QY 645 SerGluAlaLeuGluAlaLeuGluAsnGlyIleLysSerSerArgArgLeuGluAsnPhe 664
 DB 303 AGCGAGGCTTTGAGAGGCTTGGAGAAATGGAATCAAGAGCTCCCGGCGGTGAGAACTTC 362
 QY 665 CysArgAspPheGluLeuGlnLysValCysTyrLeuProLeuAsnThrPheLeuLeuArg 684
 DB 363 TGCAGAGACTTTGAGCTGAGAGAGGTGTGTACCTACCGCTCAACCTTCCCTGCGG 422
 QY 685 ProLeuHisArgLeuMetHisTyrLysGlnValIleGluLysGluCysLysHisIlePro 704
 DB 423 CCACTGACCGGCTTCACTGCACTCAAGAGAGCTCTGAGAGGCTGTGCAAAACACACCG 482
 QY 705 ProSerHisAlaAspPheArgAspCysArgAlaAlaLeuAlaGluIleThrGluMetVal 724
 DB 483 CCGAGCCAGCGCGACTTCAGAGGACTCCGAGCGCTTTGGCAGAGATCAAGGAGATGATG 542
 QY 725 AlaGlnLeuHisIleGlyThrMetIleLysMetGluAsnPheGlnLysLeuHisIleGluLys 744
 DB 543 GCACACTCCACCGTTCAGATGATGAGAGAAATTTTCCAGAAAGCTGACGAACTCAAG 602
 QY 745 LysAspLeuIleGlyIleAspAsnLeuValProGlyArgGluPheIleArgLeuGly 764

DB 603 AAGATTGATGGCATTCGATCTTGTTCCGGGAAGGAGTTCACTCCGCTGGGC 662
QY 765 SerLeuSerIyAeuSerGlylySGlyLeuGlnGlyAryMetPhePheLeuPheAAsp 784
DB 663 AGCTTAGCAAGCTCTCGGGGAAGGGGCTCAGAGGCAATGTTCTTCGTTCAAGAC 722
QY 785 ValLeuLeuTyThrSerArgGlyLeuThrAlaSerAsnGlnPheLeuValHisGlyGln 804
DB 723 GTCTCTCTAATCAACGAGCCGGGGCTGACGGCTCCAAATCAAGTTAAAGTCCAGGGCAG 782
QY 805 LeuProLeuTyGlyMetThrIleGluGluSerGluAspGlyTrpGlyValProHisCys 824
DB 783 CTCCTCCCTATATGATGAGATTGAGAGAGCAAGACAGTGGGGGGTGGCCCACTGC 842
QY 825 LeuThrLeuArgGlyGlnArgGlnSerIleIleValAlaAlaSerSerArgSerGluMet 844
DB 843 CTAACTCCCTGGGGCCAGCGGCACTCATATGATGAGCCCACTTCCTGGTCCAGATG 902
QY 845 GluLeuTrpValGluAspIleGlnMetAlaIleAspLeuAlaGluIySerSerSerPro 864
DB 903 GAGAACTGGTTAGAGCATCCAGATGGCATTGACTGGCGGAGAAGACAGACAGCCCC 962
QY 865 AlaProGluPheLeuAlaSerSerProProAspAsnIySerProAspGlyAlaThrAla 884
DB 963 GCCCTGAGATTCTGGGCAAGCAGCCCTCCGACAAACAAAGTCCCTGATGAGCAGCCGCG 1022
QY 885 AlaAspGlnGluSerGluAspAspLeuSerAlaSerArgThrSerLeuGlnArgGlnAla 904
DB 1023 GCTGACCAAGAGTCAAGAGATGACTGAGGCTTGGCC-NCATCGCTGAGAGCCCAAGCC 1081
QY 905 ProHisArgGlyAsnThrMetValHisValCysTrpHisArgAsnThrSerValIserMet 924
DB 1082 CGCACCGGGGCAACCAATGGTGCACGTGTGGTGGCAGCGCAACACACGTCCTCATG 1141
QY 925 ValAspPheSerIleAlaValGluAsnGlnLeuSerGlyAsnLeuLeuAlglyPheIyS 944
DB 1142 GTGACTTTCAGCATCGAGAGTGAAGATCACTTGTGGAAACCTGCTGAGGAATTCGAA 1201
QY 945 AsnSerAsnGlyTrpGlnIyLeuTrpValIlePheThrAsnPheCysLeuPhePheTy 964
DB 1202 AACAGCAACGGGTGGGCAAGAGCTGTGGGTGTTCACAACTTGTGCTGTTCTTCTAC 1261
QY 965 IySerHisGlnAspAsnHisProLeuAlaSerLeuProLeuGlnGlyTySerLeuThr 984
DB 1262 AATTCACACAGGACATCATCTCCCTGGCAGCTGCTGCTGGCTACTGCTCACC 1321
QY 985 IleProSerGluSerGluAsnIleGlnIyAspTyValIlePheIyLeuHisPheIySer 1004
DB 1322 ATCCCTCTGAGTCCGAGAACATCCAGAAAGACTGATGTTCAAGCTGCACTTCAAGTCC 1381
QY 1005 HisValTyTrpPheArgAlaGluSerGlyTyThrPheGlnIyArgTrpMetGluValIle 1024
DB 1382 CAGCTTACTACTTTCAGGGGGGGAAGGAGTACAGCTTCAAAAGGTGAGTGAAGTATC 1441
QY 1025 ArgSerAlaThrSerSerAlaSerArgProHisValIleuSerHisIySGluSerLeuVal 1044
DB 1442 CGCAGTGCACCAAGCTCTGCTGCGACCCCAAGTGTGAGCAAAAGAGTCTTTGAG 1501
QY 1045 Tyr 1045
DB 1502 TAT 1504

RESULT 10
ADQ63863
ID ADQ63863 standard; cDNA; 3200 BP.
XX
AC ADQ63863;
XX
DT 07-OCT-2004 (first entry)
XX
DE Novel human cDNA sequence #1024.
XX
KW 88; gene; osteopathic; neuroprotective; nootropic; antiparkinsonian;

KW cyostatic; gene therapy; diagnostic marker; morbid state; osteoporosis;
KW neurological disease; Alzheimer's disease; Parkinson's disease; dementia;
KW cancer.
XX
OS Homo sapiens.
XX
PN BP1440981-A2.
XX
PD 28-JUL-2004.
XX
PF 21-JAN-2004; 2004EP-00001196.
XX
PR 21-JAN-2003; 2003JP-00102206.
PR 09-MAY-2003; 2003JP-00131392.
XX
PA (REAS-) RES ASSOC BIOTECHNOLOGY.
PI Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
PI Yamamoto U, Isono Y, Nagai K, Irie R;
XX
DR WPI; 2004-535376/52.
XX
PT P-PSDB; ADQ6051.
XX
PT Novel 2495 cDNA, useful for treating osteoporosis, neurological diseases,
PT Alzheimer's diseases, Parkinson's diseases, dementia and various cancers.
XX
PS Claim 1; SEQ ID NO 1024; 2449pp; English.
XX
CC The invention relates to 2495 novel polynucleotides (I) and their encoded
CC polypeptides, sequences hybridizing to these nucleotides, sequences
CC encoding partial polypeptides and sequences having 70% or 90% identity to
CC the nucleotide and protein sequences. The nucleotides and polypeptides
CC are useful as diagnostic markers or therapeutic target for the diseases
CC or morbid states. They are also useful for treating osteoporosis,
CC neurological diseases, Alzheimer's diseases, Parkinson's diseases,
CC dementia and various cancers. This sequence corresponds to a nucleotide
CC sequence of the invention.
XX
SQ Sequence 3200 BP; 951 A; 659 C; 718 G; 872 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 4.01e-68 Length: 3200
Score: 1150.50 Matches: 265
Percent Similarity: 54.5% Conservative: 89
Best Local Similarity: 40.8% Mismatches: 185
Query Match: 21.1% Indels: 111
DB: Gaps: 15

US-09-555-342B-2 (1-1045) x ADQ63863 (1-3200)

QY 33 ProThrProSerGlyIyLeuValSerIleIyGlnMetLeuAspThrGlnGln 52
DB 161 CCAAGGCTCTCCGTAGATGATGATTAATAAGGCAATTTTGGATGATCCGAGAG 220
QY 53 AlaPheGluValProGlnArgAlaProGlyIyValLeuLeuAspAlaValCysAsnHis 72
DB 221 ATTTTGTGGTTGATCAAAAGCATCCGGGAAGGCAATGTTTAAACCTGAGTGCACCAT 280
QY 73 LeuAsnLeuValGluGlyAspTyThrPheGlyLeuGluPheProAspHisIySylIeThr 92
DB 281 CTAAATCTTGCTGAAGAAAGAAATTTTGGATTGAAATTTGCAACCATCTGAAATAAT 340
QY 93 ValTrpLeuAspLeuLeuIySerProIleValIyGlnIleArgArgProIyHisValVal 112
DB 341 GTTTGGCTGAGCTTTTGAAGCCATTAACAAAGCAGGTAAATAATCTTAAGAGATTGTT 400
QY 113 ValIyPheValValIyPhePheProProAspHisThrGlnLeuGlnGluGluLeuThr 132
DB 401 TTCAATTTATGATGAATTTTCCAGTGCAGCCCTGACATCTGGGGAAGAACTTACA 460
QY 133 ArgTyTrpLeuPheAlaLeuGlnValIySGlnAspLeuAlaGlnGlyArgLeuThrCysAsn 152
DB 461 AGGTATCTTTTACTCTCAATAAAGAAAGATTGGCTCTTGAGAAAGCTTCCATCAGT 520

```

QY 153 AsptHrSerAlaAlaLeuLeuLeuSerHisIleValGlnSerGluIleGlyAspPheAsp 172
Db 521 GACAACTGTACAGGCTGATGATCTCACAATCTTACATACAGAACTTGAGACTTTCAT 580
QY 173 GluAlaLeuAspArgGluHisIleValAsnValSerIleProGlnGlnAspAlaLeu 192
Db 581 GAAGAACAAGATAGAGAGCATCTGGCACAACCTGGATCTTACCAACCAAGACATGTTTA 640
QY 193 GluAspValIleValGluPheHisIleValGlnGlnIleThrProAlaGlnSerAsp 212
Db 641 GAGGCAAGATGACACTTTCATCAGAAACATTTGCGAGAGCCCGCAAGTTCCTGAC 700
QY 213 PheGlnLeuGlnIleAlaArgArgLeuGlnMetTyrGlyIleArgLeuHisProAla 232
Db 701 ATTTGCTACTTGACATAGCAAGAAAGCTGGATAGTAGGCAATCAGCCCTCACCCGCC 760
QY 233 LysAspArgGluGlyThrLysIleAsnIleValAlaAlaAsnThrGlyIleLeuValPhe 252
Db 761 AGTATGGTGAAAGGATGACATTCACCTGGCTGTGTGTCACATGGAGATCTGTGTGTA 820
QY 253 GlnGlyPheThrLysIleAsnAlaPheAsnThrAlaLysValArgLysLeuSerPheLys 272
Db 821 CGGGGAATACAAAGATCAATATCTTTAACTGGGCTAAATCCCGCAAGTTGAGTTTAAAG 880
QY 273 ArgLysArgPheLeuIleLysLeuArgProAsnAlaAsnSerAlaTyrGlnAspThrLeu 292
Db 881 AGAAAGCAATTTCTCATAACTTCAATGCCAATATCTTGCTGTGCAAGAGATACCTTG 940
QY 293 GluPheLeuMetAlaSerArgAspPheCysLysSerPheThrLysIleCysValGlnHis 312
Db 941 GAGTTCACCATGGCCAGCCGAGATGCCCTGCAAGGCTTCTGGAAGACTGTGTGGAATAC 1000
QY 313 HisAlaPhePheArgLeuPheGlnGluProLysProLysProLysProValLeuPheSer 332
Db 1001 CATCTCTTCTTCAGGCTTTCGGAAGAGCCCAATCAAGCCCAAAACCTCTCTGCGAGC 1060
QY 333 ArgGlySerSerPheArgPheSerGlyArgThrGlnLysGlnValLeuAspTyrValLys 352
Db 1061 AAGGCTTCAGCTTTCCTGATAGTGAGCAACCCAAAGCAACTTTTGAAATAGGGAGA 1120
QY 353 GlnGlyLysIleLysLysValGlnPheGlnValArgLysHis-----SerLysIleHisSer 370
Db 1121 AAAGGAGGCTGAAAGAGCTTCCCATTTGAAGAAACATTCACCATCTCAGTACATGAA 1180
QY 371 IleArgSerLeuAlaSerGlnProThrGlnLeuAsnSerGluValLeuGlnIleSerGln 390
Db 1181 ---CGACAGTGCAGGTCTCTCAACA---GACCTCTCTCTGATGTGTCAAAACAA---GTG 1231
QY 391 GlnSerThrSerLeuThrPheGlyGluGly-----Ala 401
Db 1232 GAACATTTGAGACTAGCATATGATGGTGGCTACTACCAAAATGGAATGAGTGCAGCGCA 1291
QY 402 GluSerProGlyGlyGlnSerCysArgArgGlyLysGluProLysValSerAlaGlyGlu 421
Db 1292 TCTGAGCAGTGCCTGGAGAGTAGAGAGAAATTCGATTTGAGAGTGCACATTTTGCAACT 1351
QY 422 ProGlySerHisProSerProAlaProArgArgSerProAlaGlyAsnLysGlnAlaAsp 441
Db 1352 GAGCTGAGAGCTTCCAAA-----CCAGAGCGGATCCCAATGTCATACAGTCCCAA 1405
QY 442 GlyAlaAlaSerAlaProThrGlnGluGlnGluValValLysAspArgThrGlnGln 461
Db 1406 AGCAGTTCCTCTTCCCTTT-----ATTATATGAGACCTGTCTTTAAC 1450
QY 462 SerLysProGlnPro-----ProGlnPro----- 469
Db 1451 ACTAGGCCATCTTAACCTCGATCCAGAGACATTTTTCAGAGAGAGTTCCTAAGC 1510
QY 470 -----SerThrGlySerLeuThrGlySerProHisLeuSerGluLeuSer----- 484
Db 1511 TCCTTCCAACCAACMACTGTAACTTTCTGTGTAATACATGACATATATTTCTGGGCTGACA 1570

```

```

QY 485 -----Val 485
Db 1571 AGCAAAATGCGTCCAGCAAAAGCACTTAATCAACGATGTGCCCTTATTTCTTTACA 1630
QY 486 AsnSerGlnGlyGlyValAlaProAlaAsnValThrLeuSerProAsnLeuSerProAsp 505
Db 1631 GGTGACAGAGTGGTATATATGCTCTCCAGAGTCTTTTATATGAGCAAGCAACCCAG 1690
QY 506 ThrLysGlnAlaSerProLeuIleSerProLeuLeuAsnArgIleAlaCysProArgThr 525
Db 1691 GTGCCAGATGGTCCCA----- 1708
QY 526 AspArgGluAspGluGlyArgArgLysArgPheProThrAspLysValTyrPheIleAla 545
Db 1709 -----ATTAGA 1714
QY 546 LysGluValSerThrThrGlnArgThrTyrLeuLysAspLeuGluValIleThrSerTrp 565
Db 1715 GCAGAGGAAGAAGCAAGATCCACATAGCTATGTA----- 1747
QY 566 PheGlnSerThrValSerLysGluAspAlaMetProGluAlaLeuLysSerLeuIlePhe 585
Db 1748 -----GAGCCCACTGCAATGAAAGCCAGCTGAAGAGAC----- 1780
QY 586 ProAsnPheGluProLeuHisLysPheHisThrAsnPheLeuLysGluIleGlnArg 605
Db 1781 CCAAGAAATTCAGAAATGAAGAGCTTTCAGCAAGAC---CTGCAAGTATCTCAAGAAAGCT 1837
QY 606 LeuAlaLeuTrpGluGlyArgSerAsnAlaGlnIle-----ArgAspTyrGlnArg 622
Db 1838 ATAGCCAGCACTAGCGGTAGAGAGCAATCATATAGTGTAGAAAGAGAAACCAAT 1897
QY 623 IleGlyAspValMetLeuLysAsnIleGln 632
Db 1898 TTGGAAGATGCATTTGTATGTATCAATTCAA 1927

RESULT 11
ID MAS27016 standard; cDNA; 1718 BP.
AC MAS27016;
DT 07-NOV-2001 (first entry)
XX
DE cDNA encoding novel signal transduction pathway protein, Seq ID 51.
XX
KW Neuroprotective; cytostatic; dermatological; immunosuppressive; tumour;
KW antiinflammatory; anti-HIV; antibacterial; antiinflammatory; cancer;
KW immune system disorder; rheumatoid arthritis; inflammatory condition;
KW organ transplant rejection; infection; hepatitis C; blood disorder;
KW sickle cell anaemia; hyperproliferative disorder; Gaucher's disease;
KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
KW chromosomal abnormality; Down syndrome; ischaemia; renal disorder;
KW cardiovascular; respiratory; wound healing; endocrine; Addison's disease;
KW reproductive system; gastrointestinal; liver disorder; AIDS; ss;
KW acquired immune deficiency syndrome.
XX
OS Homo sapiens.
XX
PN MO200154733-Al.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001MO-US001312.
XX
PR 31-JAN-2000; 2000US-017906SP.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.

```


CC disorders (e.g. rheumatoid arthritis), inflammatory conditions, organ
 CC transplant rejections and graft versus host disease, infectious diseases
 CC (e.g. hepatitis C), bleeding disorders, haemoglobin abnormalities and
 CC other blood-related disorders (sickle cell anaemia), myeloproliferative
 CC disorders, primary haematopoietic disorders, hyperproliferative disorders
 CC (e.g. Gaucher's disease and cancer), neurodegenerative disorders (e.g.
 CC Alzheimer's disease, Parkinson's disease), chromosomal abnormalities
 CC (Down syndrome), ischaemic injury (e.g. stroke), renal disorders (e.g.
 CC glomerulonephritis), cardiovascular disorders (e.g. arrhythmias),
 CC respiratory disorders, dermatological disorders, in wound healing,
 CC epithelial cell proliferation, endocrine disorders (e.g. Addison's
 CC disease), reproductive system disorders, gastrointestinal disorder
 CC (inflammatory disorders), liver disorders (cirrhosis), as stimulators of
 CC B-cell responsiveness to pathogens, activators of T-cells, to induce
 CC higher affinity antibodies, and as a means to induce tumour proliferation
 CC in pathologies e.g. acquired immune deficiency syndrome (AIDS). AAS26976-
 CC AAS27550 represent novel signal transduction pathway protein coding
 CC sequences and PCR primers of the invention
 CC

Alignment Scores:

Pred. No.: 1.07e-58 Length: 1718
 Score: 1006.50 Matches: 189
 Percent Similarity: 78.7% Conservative: 47
 Best Local Similarity: 63.0% Mismatches: 57
 Query Match: 18.4% Indels: 7
 Gaps: 3

US-09-555-342b-2 (1-1045) x AAS27016 (1-1718)

Qy 738 GlnlyLeuHtsgLLeuLyLeuAAspleuileGlyIleAspAenLeuValProGly 757
 Db 13 CAGAACTACAGGAGCTGCACCGGACCTGGGCAATGAAACCTATGCTCTGGC 72
 Qy 758 ArgGluPheIleArgLeuGlySerLeuSerLyLeuSerGlyLyGlyLeuGlnArg 777
 Db 73 AGGAGATTCACTCCGTGAGGGCTGCTTCACAAGCTCACCAAGAGGGCTCAGACAGG 132
 Qy 778 MetPhePheLeuPheAspValLeuLeuLyThrSerArgLyLeuThrAlaSerAsn 797
 Db 133 ATGTTTCTGTTCTCAGATATGTGCTGACACCAAGCAAGAGGTTGACAGGACCGC 192
 Qy 798 GlnPheLyLeuHtsgLLeuProLeuLyrgLyMetThrIleGlnGluSerGluAsp 817
 Db 193 CACTTCGGGATCCGGGCTCTTCCCTCCAAAGCATGCTGGTGAAGAAAGATGATAC 252
 Qy 818 GluTrpGlyValProHtsgLyLeuThrLeuArgLyGlnArgGlnSerIleIleValAla 837
 Db 253 GAGTGGTCTGTTCCACACTGTTTCACCATCAAGGGCTCAAAAACATGCTGGCA 312
 Qy 838 AlaSerSerArgSerGluMetGluLyTrpValGluAspIleGlnMetAlaIleAspLeu 857
 Db 313 GCCAGCACTCGGCTGGAAGAAAGAGTGTGACCTGAACCTCCGCGATCCAGCA 372
 Qy 858 AlaGluLySerSerSerProAlaProGluPhe-----LeuAlaSerSerPro 873
 Db 373 GCCAAGAGTGGCGGTGACAGGCCCCCTGCACTGCCAGGCCGCACTGTGTGACTGCTCC 432
 Qy 874 ProAspAnlySerProAspGlyAlaThrAlaIleAspGlnGluSerGluAspLeu 893
 Db 433 CCC-----AGATCCCCCAAGAG---GTATCTCTGAGGACGAGTCAAGAAATGATGCT 483
 Qy 894 SerIleSerArgThrSerIleuGluArgGlnAlaProHtsgLyLeuThrMetValHis 913
 Db 484 CGGGGTCTCCGACCTCCGTGAGGGGATGCGCAGCACCAGGCAACCAACATATCAC 543
 Qy 914 ValCytrpHtsgAenThrSerValSerMetValAspPheSerIleAlaValGluAsn 933
 Db 544 GTGTGCTGTGATCCGGAACCAAGCGTGTCCAGGGCAACCAACATGACGCTGTGAAGAC 603
 Qy 934 GlnLeuSerGlyAnlyLeuLeuArgLyPhePheLyAsnSerAsnGlyTrpGlnLyLeuTrp 953
 Db 604 CAGCTTCAGGATATCTGCTAAGAAAGTTCAAAAAGTCAATGCTGGCAGAACTCTGG 663

Qy 954 ValValPheThrAspPheCyLeuPhePheTyrrysSerHtsgIleAspAnlyProLeu 973
 Db 664 GTGCTTTTACCACTTCTGTTTCTTCTTCAAAAATCTACAGATGACTTACCTG 723
 Qy 974 AlaSerLeuProLeuLeuGlyTyrrysLeuThrIleProSerGluSerGluAsnIleGln 993
 Db 724 GCCAGCTCCCGGCTGCTGGGTACAGGATGAGCATCCCGAGGAGCGCATGCAATCAC 783
 Qy 994 LyAspTrpValPheLyLeuHtsgPheLySerHtsgValTyrrysPheAlaGluSer 1013
 Db 784 AAAGACTATGTTTCAAGCTCCAGTTCAAAATCCACGCTTACTTCCGGGCTGAAGAC 843
 Qy 1014 GluTrpThrPheGluArgTrpMetGluValIleArgSerAlaThrSerSerAlaSerArg 1033
 Db 844 AAGTACATTTGAAGGTGATGAGTATCCAGGGGCGACAGCTCAGCCGGAGG 903

RESULT 12
 ADB93194
 ID ADB93194 standard; cDNA; 1718 BP.
 XX AC
 XX ADB93194;
 XX DT 04-DEC-2003 (first entry)
 XX DE
 XX Human cDNA encoding a novel protein #41.
 KW 88; gene; human; autoimmune disease; Parkinson's disease; slliosis;
 KW gastrointestinal disease; atherosclerosis; haemophilia; thrombocytopenia;
 KW immunosuppressive agent; adjuvant; enhance immune response;
 KW higher affinity antibody induction;
 KW increased serum immunoglobulin concentration.
 OS Homo sapiens.
 PN US2002168711-A1.
 XX PD 14-NOV-2002.
 XX PF 17-JAN-2001; 2001US-00764868.
 XX PR 31-JAN-2000; 2000US-0179065P.
 PR 04-FEB-2000; 2000US-0180628P.
 PR 28-JUN-2000; 2000US-0214886P.
 PR 07-JUL-2000; 2000US-0216647P.
 PR 07-JUL-2000; 2000US-0216880P.
 PR 11-JUL-2000; 2000US-0217487P.
 PR 11-JUL-2000; 2000US-0217496P.
 PR 14-JUL-2000; 2000US-0218290P.
 PR 26-JUL-2000; 2000US-0220963P.
 PR 26-JUL-2000; 2000US-0220964P.
 PR 14-AUG-2000; 2000US-0224518P.
 PR 14-AUG-2000; 2000US-0224519P.
 PR 14-AUG-2000; 2000US-0225267P.
 PR 14-AUG-2000; 2000US-0225268P.
 PR 14-AUG-2000; 2000US-0225269P.
 PR 14-AUG-2000; 2000US-0225447P.
 PR 14-AUG-2000; 2000US-0225757P.
 PR 14-AUG-2000; 2000US-0225758P.
 PR 22-AUG-2000; 2000US-0226868P.
 PR 30-AUG-2000; 2000US-0228924P.
 PR 01-SEP-2000; 2000US-0229287P.
 PR 01-SEP-2000; 2000US-0229343P.
 PR 01-SEP-2000; 2000US-0229344P.
 PR 01-SEP-2000; 2000US-0229345P.
 PR 05-SEP-2000; 2000US-0229509P.
 PR 05-SEP-2000; 2000US-0229513P.
 PR 08-SEP-2000; 2000US-0231413P.
 PR 21-SEP-2000; 2000US-0234223P.
 PR 21-SEP-2000; 2000US-0234274P.
 PR 25-SEP-2000; 2000US-0234997P.
 PR 27-SEP-2000; 2000US-0235834P.
 PR 29-SEP-2000; 2000US-0236332P.

PR 29-SEP-2000; 2000US-0236367P.
 PR 29-SEP-2000; 2000US-0236368P.
 PR 29-SEP-2000; 2000US-0236369P.
 PR 29-SEP-2000; 2000US-0236370P.
 PR 02-OCT-2000; 2000US-0236802P.
 PR 02-OCT-2000; 2000US-0237037P.
 PR 02-OCT-2000; 2000US-0237038P.
 PR 02-OCT-2000; 2000US-0237039P.
 PR 02-OCT-2000; 2000US-0237040P.
 PR 13-OCT-2000; 2000US-0239935P.
 PR 20-OCT-2000; 2000US-0240960P.
 PR 20-OCT-2000; 2000US-0241785P.
 PR 20-OCT-2000; 2000US-0241809P.
 PR 01-NOV-2000; 2000US-0244617P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 08-DEC-2000; 2000US-0251865P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 XX
 PA (ROSEN/ ROSEN C A.
 PA (RUBEN/ RUBEN S M.
 PA (BARA/ BARASH S C.
 PI Rosen CA, Ruben SM, Barash SC;
 DR WPI; 2003-719985/68.
 XX P-PSDB; ADB93807.
 PT New isolated polypeptide useful for diagnosing and treating
 PT immunosuppressive conditions such as autoimmune disease and Parkinson's
 PT disease.
 XX
 PS Claim 3; SEQ ID NO 51; 345dp; English.
 XX
 CC The invention relates to an isolated polypeptide. The polypeptide is
 CC useful for diagnosing a pathological condition or a susceptibility to a
 CC pathological condition in a subject, by determining the presence or
 CC amount of expression of the polypeptide in a biological sample and
 CC diagnosing a pathological condition or a susceptibility to a pathological
 CC condition based on the presence or amount of expression of the
 CC polypeptide. The polypeptide is also useful for identifying a binding
 CC partner to the polypeptide, which involves contacting the polypeptide
 CC with a binding partner and determining whether the binding partner
 CC effects an activity of the polypeptide. The polypeptide or the nucleic
 CC acid encoding the polypeptide is useful for preventing, treating, or
 CC ameliorating a medical condition, which involves administering the
 CC polypeptide or the nucleic acid to a mammalian subject. The nucleic acid
 CC is useful for diagnosing a pathological condition or a susceptibility to
 CC a pathological condition in a subject, which involves determining the
 CC presence or absence of a mutation in the nucleic acid, and diagnosing a
 CC pathological condition or susceptibility to a pathological condition
 CC based on the presence or absence of the mutation. The polypeptide, the
 CC nucleic acid and an antibody to the polypeptide are useful for treating
 CC autoimmune disease, Parkinson's disease, slliosis, gastrointestinal
 CC disease, atherosclerosis, haemophilia, thrombocytopenia. The polypeptide,
 CC the nucleic acid and the antibody are useful as immunosuppressive agents,
 CC as adjuvants to enhance immune responses, and as agents to induce higher
 CC affinity antibodies and increase serum immunoglobulin concentrations. The
 CC present sequence represents cDNA encoding a novel human protein. Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification but was obtained in electronic format direct from USPTO at
 CC seqdata.uspto.gov/sequence.html?DocID=20020168711.
 CC
 XX
 SO Sequence 1718 BP; 428 A; 500 C; 476 G; 314 T; 0 U; 0 Other;

Alignment Scores:

Pred. NO.: 1.07e-58 Length: 1718
 Score: 1006.50 Matches: 189
 Percent Similarity: 78.7% Conservative: 47
 Best Local Similarity: 63.0% Mismatches: 57
 Query Match: 18.4% Indels: 7
 Gaps: 3

US-09-555-342b-2 (1-1045) x ADB93194 (1-1718)
 QY 738 GlnYLeuNIeGluLeuYsYAspLeuIleGlyIAspAsnLeuValProGly 757
 Db 13 CAGAGCTAACGAGCTGCGAGCGGAGCTGTGGGAGATAGAACTCATGTCTCTGGC 72
 QY 758 ArgGluPheIleArgLeuGlySerLeuSerIlyAsGlyLeuGlnArg 777
 Db 73 AGGAGATTCATCCGTTAGGGCTGCTTCACAAAGCTCACAAGAAAGGCGCTGCAGAGAG 132
 QY 778 MetPhePheLeuPheAsnAspValLeuLeuYrThSerArgGlyLeuThrAlaSerAsn 797
 Db 133 AGTTTTTCTGTCTTCAGATATGTTGTCGTACACAAGAAAGAGTTCAGGGAGCCAGC 192
 QY 798 GlnPheValIleGlyGlnLeuProLeuYrGlyMetThrIleGluGlnSerGluAsp 817
 Db 193 CACTTCGGATCCGGGCTCTCTCTCCCTCCAAAGCATCTGTGTGAAGAAATGATTAAC 252
 QY 818 GluTrpGlyValProHisCysLeuThrIleAsArgGlyGlnArgGlnSerIleIleValAla 837
 Db 253 GAGTGTCTGTTCACACATGTTTCACCATCTACCGCGCTCAGAAACAAATCGTGTGGCA 312
 QY 838 AlaSerArgSerGluMetGluYrTrpValGluAspIleGlnMetAlaIleAspLeu 857
 Db 313 GCCAGCAGCTCGCTGAGAAAGAGTGAATGATGACTGAACCTCCGAGTCCAAAGCA 372
 QY 858 AlaGluYsSerSerProAlaProGluPhe-----LeuAlaSerSerPro 873
 Db 373 GCCAAGATGGCGGTACACGCGCTTCGATCCAGCGCCGACCTGTGTGCACTGCTCC 432
 QY 874 ProAspAsnIlySerProAspGluAlaThrAlaAlaAspGlnIlySerGluAspAspLeu 893
 Db 433 CCC-----AGATCCCCCAAGCAG--GATCTCTGAGACAGGAGTCAGAAATGATGCT 483
 QY 894 SerAlaSerArgThrSerLeuGluArgGlnAlaProHisArgGlyAsnThrMetValHis 913
 Db 484 CGGGGTGTCCGACGCTCTCTGGAGGGGATGGCCAGACCGGGCCAAACACCAATGAC 543
 QY 914 ValCysTrpHisArgAsnThrSerValSerMetValAspPheSerIleAlaValGluAsn 933
 Db 544 GTGTGTGTGTACCGGAACACACGCGTGTCCAGGGCAGACCAAGTCCAGTGTCCAGAAC 603
 QY 934 GlnLeuSerGlyAsnLeuLeuArgIlyAspAsnSerAsnGlyTrpGlnIlyLeuTrp 953
 Db 604 CAGCTTCAGATATGCTGCTAAGAAAGTCAAAAACAGTCATGCTGGCAGAAAGCTCTGG 663
 QY 954 ValValPheThrAsnPheCysLeuPhePheTrpIlySerHisGlnAspAsnHisProLeu 973
 Db 664 GTGCTTTTACCAACTTGTGTGTCTTCTCAAAAATCATCAGATGATGATGCCACTG 723
 QY 974 AlaSerLeuProLeuLeuGlyTrpSerLeuThrIleProSerGluSerGluAsnIleGln 993
 Db 724 GCCAGCTCCCGCTGTGGGCTTACAGCTGAGATGCCACGAGAGCCCATGSCATACAC 783
 QY 994 LysAspTrpValPheIlyLeuHisPheIlySerHisValIlyTrpPheArgAlaGluSer 1013
 Db 784 AAAGACTATGTTTCAAGCTCCAGTCAATATCCCAAGTCTACTTCTCCGGCTGAGAGC 843
 QY 1014 GluTrpThrPheGluArgTrpMetGluValIleArgSerAlaThrSerSerAlaSerArg 1033
 Db 844 AAGTACACATTTGAAAGGTGAGTGAAGGTGATTCAGAGGGGCCAGAGCTCAGCGGAGG 903
 RESULT 13
 ACD92278/c
 ID ACD92278 standard; cDNA; 582 BP.
 XX
 AC ACD92278;
 XX
 DT 23-SEP-2003 (first entry)
 XX
 DE Human colon cancer cell expressed cDNA #690.
 KW Open reading frame detection; genome sequencing; colon cancer;

KM breast cancer; population genome analysis; genetic shift; cancer;
 KM antibiotic resistance; antibiotic non-tolerance; congenital disease;
 KM agriculture; food crop genome; resistance gene; retrovirus;
 KM influenza virus; eukaryotic pathogen detection; trypanosome; Plasmodium;
 KM gene; ss.
 XX Homo sapiens.
 XX US2002155438-A1.
 XX 24-OCT-2002.
 XX 27-SEP-1999; 99US-00406117.
 XX 20-NOV-1998; 98US-00196716.
 XX (SIMP/) SIMPSON A J G.
 XX (NETO/) NETO E D.
 XX (BREN/) BRENTANI R R.
 XX Simpson AJG, Neto ED, Brentani RR;
 DR WPI; 2003-182626/18.
 PT Determining open reading frames of genome of an organism e.g. a human
 PT suffering from cancer involves use of single oligonucleotide primer at
 PT low stringency for preparing single-stranded cDNA from mRNA of
 PT individual.
 XX Example 9; Page 105; 959pp; English.
 XX The invention describes a method of determining open reading frames in
 CC the genome of organism comprising contacting mRNA from cell of organism
 CC with a single oligonucleotide primer (I) at low stringency, preparing
 CC single-stranded cDNA by reverse transcribing mRNA with (I), amplifying
 CC cDNA, sequencing the product, and repeating the contacting, preparing
 CC and amplifying steps with different primers and sequencing resulting
 CC nucleic acids. The method is useful for: determining that a known
 CC nucleotide sequence from a genome of an organism corresponds to a
 CC nucleotide sequence from an open reading frame; for preparing a contig,
 CC nucleic acid molecule from a genome of an organism; and for sequencing
 CC all or part of a genome of an organism. mRNA is obtained from mammalian
 CC or human cell which is associated with a pathological condition e.g. a
 CC colon cancer or breast cancer cell. The method is useful for analyses of
 CC populations of subjects and can be used to carry out genetic analyses of
 CC large or small populations. Further, it can be used to study living
 CC systems to determine if, e.g. there have been genetic shifts which render
 CC an individual or population more or less likely to be afflicted with
 CC diseases such as cancer, to determine antibiotic resistance or non-
 CC tolerance, and so forth. The method can also be used in the study of
 CC congenital diseases, and the risk of affliction to a foetus, as well as
 CC the study of whether the conditions are likely to be passed to offspring
 CC through ova or sperm. The analyses for pathological conditions can be
 CC carried out in all animals, plants, birds, fish, etc. Using this method,
 CC in the area of agriculture, for example the genomes of food crops can be
 CC studied to determine if resistance genes are present, defects in plant
 CC genomes can also be studied in this way. Similarly, the method permits
 CC determination of the pathogens which integrate into the genome, such as
 CC retroviruses and other integrating viruses such as influenza virus, have
 CC undergone shifts or mutations, which may require different approaches to
 CC therapy. This method is also applied to eukaryotic pathogens, such as
 CC trypanosomes, different types of Plasmodium, etc. The method essentially
 CC eliminates sequencing of non-coding portions. This sequence represents a
 CC polynucleotide isolated from human colon cancer cell cDNA library
 XX
 XX Sequence 582 BP; 109 A; 147 C; 185 G; 140 T; 0 U; 1 Other;

DB: 10 Gaps: 0
 US-09-555-342B-2 (1-1045) x ACD92278 (1-582)
 QY 859 GluysSerSerSerProAlaProGluPheLeuAlaSerSerProProAspAsnLysSer 878
 Db 579 GAAAGAGACAGACAGCCCCGCCCTTGAGATTCTGTGCGCAGACGCCCTCGACAAAGTCC 520
 QY 879 ProAspGluAlaThrAlaAlaAspGlnGluSerGluAspAspLeuSerAlaSerArgThr 898
 Db 519 CCGATATAGCCACCGCGGCTGACACAGAGTCAAGATGATGACGCGCTCGGACCA 460
 QY 899 SerLeuGluArgGlnAlaProHisArgGlyAsnThrMetValHisValCysTrpHisArg 918
 Db 459 TCGCTGAGCGCCAGCGCCCGCCAGCCGCGCAACCAATGTGTGCTGTGCAACGC 400
 QY 919 AsnThrSerValSerMetValAspPheSerIleAlaValGluAsnGlnLeuSerGlyAsn 938
 Db 399 AACACACAGCGTCTCCATGTGAGTGAATCTCAGCATCGCAGTGAATCAGTTGTGAAAC 340
 QY 939 LeuLeuArgLysPheLysAsnSerAsnGlyTrpGlnLysLeuTrpValAlaPheThrAsn 958
 Db 339 CTGCTGAGGAATTCATCAAAAACAGCAACGGGTGGCAAGCTGTGGTGTTCACAAAC 280
 QY 959 PheCysLeuPhePheTyrlsSerHisGlnAspAsnHisProLeuAlaSerLeuProLeu 978
 Db 279 TTCGCTGCTTCTTCTTCAAAATCACACAGACCAATATCCCTTGCCAGCTGCTCTG 220
 QY 979 LeuGlyTyrSerLeuThrIleProSerGluSerGluAsnIleGlnLysAspTyrValPhe 998
 Db 219 CTCGGTACTGCTGCACCATCCCTCTGAGTCCGAGAACATCCAGAAAGCTAGCTGTTCC 160
 QY 999 LysLeuHisPheLysSerHisValTyrTyrPheArgAlaGluSerGlyTyrThrPheGlu 1018
 Db 159 AAGCTGACCTTCAGTCCAGTCTTACTTCTGAGGCGGAAACCGTACACCTTGAA 100
 QY 1019 ArgTrpMetGluValIleAspSerAlaThrSerSerAlaSerArgProHisValLeuSer 1038
 Db 99 AGGTGATGAGAGTATCCGACGTGCCACCACTCTGCTGCGACCCACGTGTGAGC 40
 QY 1039 HisLysGluSerLeuValTyr 1045
 Db 39 CACAAAGAGTCTCTGTGTAT 19
 RESULT 14
 ACD94832/c
 ID ACD94832 standard; cDNA; 582 BP.
 AC ACD94832;
 XX 23-SEP-2003 (first entry)
 DT Human colon cancer cell expressed cDNA #3244.
 XX
 XX Open reading frame detection; genome sequencing; colon cancer;
 KM breast cancer; population genome analysis; genetic shift; cancer;
 KM antibiotic resistance; antibiotic non-tolerance; congenital disease;
 KM agriculture; food crop genome; resistance gene; retrovirus;
 KM influenza virus; eukaryotic pathogen detection; trypanosome; Plasmodium;
 KM gene; ss.
 XX Homo sapiens.
 XX US2002155438-A1.
 XX 24-OCT-2002.
 XX 27-SEP-1999; 99US-00406117.
 XX 20-NOV-1998; 98US-00196716.
 XX (SIMP/) SIMPSON A J G.
 XX (NETO/) NETO E D.

Alignment Scores:
 Pred. No.: 2.01e-56 Length: 582
 Score: 964.00 Matches: 183
 Percent Similarity: 98.4% Conservative: 1
 Best Local Similarity: 97.9% Mismatches: 3
 Query Match: 17.6% Indels: 0


```

PA (BREN//BRENTANI R. R.
XX
PI Simpson AUG, Neto ED, Brentani RR;
XX
XX WPI; 2003-182626/18.
XX
PT Determining open reading frames of genome of an organism e.g. a human
PT suffering from cancer involves use of single oligonucleotide primer at
PT low stringency for preparing single-stranded cDNA from mRNA of
PT individual.
PS
PS Example 9; Page 477; 9599p; English.
XX
XX The invention describes a method of determining open reading frames in
CC the genome of organism, comprising contacting mRNA from cell of organism
CC with a single oligonucleotide primer (I) at low stringency, preparing
CC single-stranded cDNA by reverse transcribing mRNA with (I), amplifying
CC cDNA, sequencing the product, and repeating the contacting, preparing
CC and amplifying steps with different primers and sequencing resulting
CC nucleic acids. The method is useful for: determining that a known
CC nucleotide sequence from a genome of an organism corresponds to a
CC nucleotide sequence of an open reading frame for preparing a contig,
CC nucleic acid molecule from a genome of an organism; and for sequencing
CC all or part of a genome of an organism. mRNA is obtained from mammalian
CC or human cell which is associated with a pathological condition e.g. a
CC colon cancer or breast cancer cell. The method is useful for analyses of
CC populations of subjects and can be used to carry out genetic analyses of
CC large or small populations. Further, it can be used to study living
CC systems to determine if, e.g. there have been genetic shifts which render
CC an individual or population more or less likely to be afflicted with
CC diseases such as cancer, to determine antibiotic resistance or non-
CC tolerance, and so forth. The method can also be used in the study of
CC congenital diseases, and the risk of affliction to a fetus, as well as
CC the study of whether the conditions are likely to be passed to offspring
CC through ova or sperm. The analyses for pathological conditions can be
CC carried out in all animals, plants, birds, fish, etc. Using this method,
CC in the area of agriculture, for example the genomes of food crops can be
CC studied to determine if resistance genes are present, defects in plant
CC genomes can also be studied in this way. Similarly, the method permits
CC determination of the pathogens which integrate into the genome, such as
CC retroviruses and other integrating viruses such as influenza virus, have
CC undergone shifts or mutations, which may require different approaches to
CC therapy. This method is also applied to eukaryotic pathogens, such as
CC trypanosomes, different types of Plasmodium, etc. The method essentially
CC eliminates sequencing of non-coding portions. This sequence represents a
CC polynucleotide isolated from human colon cancer cell cDNA library
XX
XX
SO Sequence 582 BP; 109 A; 147 C; 185 G; 140 T; 0 U; 1 Other;

Alignment Scores:
Pred. No.: 2,01e-56 Length: 582
Score: 964.00 Matches: 183
Percent Similarity: 98.4% Conservative: 1
Best Local Similarity: 97.9% Mismatches: 3
Query Match: 17.6% Indels: 0
DB: 10 Gaps: 0

US-09-555-342B-2 (1-1045) x ACD94832 (1-582)
OY 859 GLuYsSerSerSerProAlaProGluPheLeuAlaSerSerProProAspAntySer 878
Dy :::::
Db 579 GAAGAAGACAGCAGCCGCCCTTAGTTCCTGCGCAGCAGCCCCCTTAGACAAAGTCC 520
OY 879 ProAbgGluAlaEthAlaAlaAbpGlnGluSerGluAspAspLeuSerAlaSerArgTr 898
Db 519 CCTGATGAAGCCACCGCGCTAGCAGCGAGTCAAGANAGACCTGAGCGCTCGCCACA 460
OY 899 SerLeuGluArgGluAlaProHisArgGlyAsnThrMetValHisAlaCysTrpHisArg 918
Db 459 TCGCTGAGAGCCGACAGCCCGACCGCGGACACACATGTCACGTGTCTGCAACCGC 400
OY 919 AnthrSerValSerMetValAspPheSerIleAlaValGluAsnGluLeuSerGlyAsn 938

```

Dd	399	ANACNACGCGCTCCATCGTGGAGCTTTCAGCATCCGACGCGAGGANTCAGTTGCTGGAAAC	340
Qy	939	LeuLeuArlGlySPheLySPaenSeSerAemGlyTTPGlnLyLeuTPValPheThrAen	958
Dd	339	CTGCTGAGGAATTCATAAAACAGCAACGGGCGGACGAAGCTGTGGTGTTCACAAAC	280
Qy	959	PheCySPhePhePheTyTLyLeSeSerHiGlnAapAenHiSProLeuAlaSeSerLeuProLeu	978
Dd	219	TTTCGCGCTGTTCTTTCACAAATCACACCGAGCAATCATCCCTTGGCAGCCTGCTCTG	220
Qy	979	LeuGlyTLySerLeuThrLileProSeGluSerGluAenLileGlnLySPaTyTLyPhe	998
Dd	219	CTCGGCTACTGCTCACCATCCCTCTGAGTCCAGAAATCCAGAAAGATACGTTCTTC	160
Qy	999	LyLeuHiSPheLySeSerHiValTLyTLyPheArlAglIuSerGluTLyThrPheGlu	1018
Dd	159	AAGCTGCATTCAGAGTCCACGCTTACTACTTCAGGGGCGAAGCGAGTACACGTTGAA	100
Qy	1019	ArgTPMeGluValLileArgSerAlaThrSeSerAlaSeArGProHiSPheValLeuSer	1038
Dd	99	AGGTGGANTGAAAGTGATCCGACAGTCCACCAAGCTCTGCTTGGCAGCCCACTGTGTGAGC	40
Qy	1039	HisLySGluSerLeuValTLyTyr	1045
Dd	39	CACAAAGAGTCTCTGTGTAT	19
RESULT 15			
ABL03867	ID	ABL03867 standard; cDNA; 1747 BP.	
XX	XX	ABL03867;	
AC	XX		
XX	XX	26-MAR-2002 (first entry)	
DT	XX		
DE	XX	Drosophila melanogaster expressed polynucleotide SEQ ID NO 6083.	
XX	XX		
KM	XX	Drosophila; developmental biology; cell signalling; insecticide;	
KW	XX	pharmaceutical; gene; ss.	
XX	XX		
OS	XX	Drosophila melanogaster.	
PN	XX	WO200171042-A2.	
PD	XX		
XX	XX	27-SEP-2001.	
XX	XX		
PF	XX	23-MAR-2001; 2001WO-US009231.	
XX	XX		
PR	XX	23-MAR-2000; 2000US-0191637P.	
PR	XX	11-JUL-2000; 2000US-00614150.	
PA	XX	(PEKE) PE CORP NY.	
XX	XX		
PI	XX	Venter JC, Adams M, Li PWD, Myers EW;	
XX	XX		
DR	XX	WPI; 2001-656860/75.	
XX	XX	P-PSDB; ABB59764.	
PT	XX		
XX	XX	New isolated nucleic acid detection reagent for detecting 1000 or more	
XX	XX	genes from Drosophila and for elucidating cell signaling and cell-cell	
XX	XX	interactions.	
PS	XX	Claim 1; SEQ ID NO 6083; 21pp + Sequence Listing; English.	
CC	CC	The invention relates to an isolated nucleic acid detection reagent	
CC	CC	capable of detecting 1000 or more genes from Drosophila. The invention is	
CC	CC	useful in developmental biology and in elucidating cell signalling and	
CC	CC	cell-cell interactions in higher eukaryotes for the development of	
CC	CC	insecticides, therapeutics and pharmaceutical drugs. The invention	
CC	CC	discloses genomic DNA sequences (ABJ16176-ABJ30511), expressed DNA	
CC	CC	sequences (ABJ01840-ABJ16175) and the encoded proteins (ABB57737-	
CC	CC	ABB572072). The sequence data for this patent did not form part of the	
CC	CC	printed specification, but was obtained in electronic format directly	
CC	CC	from Wipo at ftp:wipo.int/pub/published_pct_sequences	

XX Sequence 1747 BP: 431 A; 492 C; 461 G; 363 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 2,26e-52 Length: 1747
 Score: 913.50 Matches: 217
 Percent Similarity: 54.0% Conserved: 110
 Best Local Similarity: 35.8% Mismatches: 207
 Query Match: 16.7% Indels: 73
 DB: 4 Gaps: 14
 US-09-555-342B-2 (1-1045) x ABL03867 (1-1747)
 QY 423 GlyserrHisProSerProAlaProArgSerProAlaGlyAenlyGlnAlaAspGly 442
 DB 87 GGATCGGATAAAGCTTTCCTCACTATCCCGATCTAGGAGGAAACA----- 131
 QY 443 AlaAlaSerAlaProThrGlnGlnGlnGlnValValLysAspArgThrGlnGlnSer 462
 DB 132 -----TATGACGTAAACAGGAGAAATCCAGACGCT 164
 QY 463 LysProGlnProGlnProSerThrGlySerLeuThrGlySerProHisLeuSerGlu 482
 DB 165 AAGAGACAG-----GAGCTAGCCACAGCGCTGCGCACCGCTCG----- 203
 QY 483 LeuSerValAsnSerGlnGlyValAlaProAlaAsn---ValThrLeuSerProAsn 501
 DB 204 -----CACCGAATGGAGCGGAAACGGAATGGGACATACCTTAGC----- 245
 QY 502 LeuSerProAspThrLysGlnAlaSerProLeuLeuSerProLeuLeuAsnAspGlnAla 521
 DB 245 ----- 245
 QY 522 CysProArgThrAspAspGluAspGlnGlyArgArgLysPheProThrAspLysAla 541
 DB 246 ---ACACAAACAGATATTGAGGCGAAGTGAAGAAAGCGGAATGGCCACCGAGCCGAC 302
 QY 542 TyrPheIleAlaLysGlnValSerThrThrGlnArgThrTyrLeuLysAspLeuGlnVal 561
 DB 303 TACTTTTGGCCAAAGAGAGCTGTGATGACGAGACGTACGTAAGAAAGATCTAGATGTG 362
 QY 562 IleThrSerTrpPheGlnSerThrValSerLysGlnAspAlaMetProGlnAlaLeuLys 581
 DB 363 CTGAACACACCTTCCGACAGCTGTGCTCGGGAGATG-----GACACAGCTGACAG 416
 QY 582 SerLeuIlePheProAsnPheGlnProLeuHisLysPheHisThrAsnPheLeuLysGlu 601
 DB 417 CCGCTG---TTCGAGCTGCTGACCTCCCTGGCCGACGACCAATCTCTCTGCGGACAC 473
 QY 602 IleGlnGlnAlaGlnAlaLeuTyrGlnGlyArgSerAsnAlaGlnIleArgAspTyrGln 621
 DB 474 ATTCAAACACCCGATGCTGTCAGTGGAGGAGGAGAGGAGC-----CAGCAAGCCGAC 524
 QY 622 ArgIleGlyAspValMetLeuLysAsnIleGlnGlyMetLysHisLysLeuAlaHisLeu 641
 DB 525 CGCATCGGAGACGTCATGATAGAGCATGGCCCGCTGCGCCATCTACGATGAGTACGTG 584
 QY 642 TrpLysHisSerGlnAlaLeuGlnAlaLeuGlnLysGlnGlyLysSerSerArgArgLeu 661
 DB 585 CAGAGCGACCTGACATCTTGCACTGATGAACGACATGTACGAAGGAGGAGAACGCTTC 644
 QY 662 GlnAsnPheCysAspAspPheGlnGlnGlnLysValCysTyrLeuProLeuAsnThrPhe 681
 DB 645 CGTCAGAGGTACAGAGGATTGAGCAGCAAAAGGTTTCTATCTAACCATGAGGAGAACTT 704
 QY 682 LeuLeuArgProLeuHisArgLeuMetHisTyrLysGlnValIleGlnAlaArgLysCysVal 701
 DB 705 CTACTGAAGCCCTTACCGGCTGTGCTGACATACCACTGATCTTAGACGGGCTCTGCGAC 764
 QY 702 HisHisProSerHisAlaAspPheArgAspCysArgAlaAlaLeuAlaGlnIleThr 721
 DB 765 TACTATGGGAGAGGAGCATATGCACTATGCGGATGCTATGCGCGTGCACCACTTGCTGCT 824

QY 722 GlnMetValAlaGlnLeuHisGlyThrMetIleLysMetGlnLysPheGlnLysLeuHis 741
 DB 825 CGCAGCACCAAGGATATTAGTCCAGACTCCGAGCTCTGCAAACTTTGAGAGCTGTGT 884
 QY 742 GlnLeuLysLysAspLeuIleGlyIleAspAsnLeuValValProGlyArgGlnPheIle 761
 DB 885 GAATCGCAAGCGCAC---ATCAACTTGACAGACAGTGGTCTGACGCCCTCCCGCTTATC 941
 QY 762 ArgLeuLysSerLeuSerLysLeuSerGlyLysGlyLeuGlnGlnArgMetPhePheLeu 781
 DB 942 CGCCAGGAGATGCTTGTGAAGACCTCCAGCGCGGTGTGACAGAGAGATTTCTTTG 1001
 QY 782 PheAsnAspValLeuLeuTyrThrSerArgGlyLeuThrAlaSerAsnGlnPheLysVal 801
 DB 1002 TTCTCCGACCTGCTGCTCTTACGTTGCAAGTCC---CCGCTGACGACAGCTTCCGAT 1058
 QY 802 HisGlyLysLeuProLeuTyrGlyMetThrIleGlnGlnSerGlnAspGlnTyrGlyVal 821
 DB 1059 TTAGGCCACGTTCTGCGCTTCACTATCTACACCGAGAACCCGAGCAC----- 1106
 QY 822 ProHisCysLeuThrLeuArgGlyGlnArgGlnSerIleIleValAlaAlaSer----- 839
 DB 1107 ---AACACTTCTTCATCTTCGGGACAGTGGCCATCAGTACAGACAAAGATGACAC 1163
 QY 840 -----SerArgSerGlnMetGlnLysTyrValGlnLysPheIleGlnMetAlaIleAspLeu 857
 DB 1164 CAATCCCAAAATCTTCTTCGACAAAGCACCCAGTAAATGATATTCCACTTATTTTCA 1223
 QY 858 AlaGlnLysSerSerSerProAlaProGlnPheLeuAlaSerSerProProAsnLys 877
 DB 1224 GGCTCTTC-GAAGAGGCGCTGAGCT-----ATTGCTGTGACGAAAGCAACAG 1276
 QY 878 SerProAspGlnAlaThrAlaAlaAspGlnGlnSerLys-----AspAspLeuSer 894
 DB 1277 CAGCTTAAACAGCAGTGTGAACGGCGGCGGTCCGCTGACTTACGAGACGAGAACTTGA 1336
 QY 895 AlaSerArgThrSerLeuGlnArgGlnAla-ProHisArgGlyAsnThrMetValHisVal 914
 DB 1337 GCTGCACCAACAGCAGCAAAAGAGAGCAGCCATCCCGAAAGCAATCGGCTCGACAGT 1396
 QY 914 LysTrpHisArgAsnThrSerValSerMetValAspPheSerIleAlaValGlnAsnGln 934
 DB 1397 CTGTGGCATCTGTGGCCACCGGACGTAAGGACTGCGGATCACCTGATACCCCGGACACCA 1456
 QY 934 PheSerGlyAsnLeuLeuArgLysPheLysAsnSerAsnGlyTyrGlnLysLeuTyrVal 954
 DB 1457 GTTGTCCGATTATCTGTGCGAAAGTTCAAGAAATAGCTCCGCTGCGGAGAACTTGTGGT 1516
 QY 954 ValPheThrAsnPheCysLeuPhePheTyrLysSerHisGlnAspAsnHisProLeuAla 974
 DB 1517 GGTGTCAAGCTTGTGTGTCTGTACTTTACMAAGGCTTACAGAGCAGATTCGCACTGGC 1576
 QY 974 AserLeuProLeuLeuGlyTyrSerLeuThrIleProSerGlnSerGlnAsnIleGlnLys 994
 DB 1577 TAGCTCCCGCTATTTGGTTTACAGCGTGGGCTCCCTGTCACCAAGATGCCGTTCAAGA 1636
 QY 994 AspTyrValPheLysLeuHisPheLysSerHisValTyrTyrPheArgAlaGlnSerGln 1014
 DB 1637 GGAATGTGCTTCAAGCTTTCCTTAAAGAACACAGTCTACTTCTTCGCGCGGAAAGTGC 1696
 QY 1014 uTyrThrPheGlnLys 1019
 DB 1697 GCACACCTTACAAAGG 1712

Search completed: May 20, 2006, 03:10:22
 Job time : 1574 secs

This Page Blank (uspto)

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

Om protein - nucleic search, using frame_p2n model

Run on: May 20, 2006, 01:45:51 ; Search time 11213 Seconds
(without alignments)
7817.137 Million cell updates/sec

Title: US-09-555-342B-2
Perfect score: 5463
Sequence: 1 MGELQRPRTGSRGAPENS.....SATSSASPHYLSHESLVY 1045

Scoring table:
BIOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 48236798 seqs, 27959665780 residues

Total number of hits satisfying chosen parameters: 96473596

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+g2n.model -DEV=x1h
-Q=/abs/ABSMBB.spool/US09555342/runat_19052006_121457_13216/app.query.fasta_1
-DB=EST -QMT=fastcap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOFCR=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blowum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=dbs07
-USRR=US09555342.0CGN.1.1_6323.0runat_19052006_121457_13216 -NCPU=6 -ICPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-MAIN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST: *
1: gb_est1.*
2: gb_est3.*
3: gb_est4.*
4: gb_est5.*
5: gb_est6.*
6: gb_est7.*
7: gb_est8.*
8: gb_est9.*
9: gb_est10.*
10: gb_est11.*
11: gb_est12.*
12: gb_est13.*
13: gb_est14.*
14: gb_est15.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5463	100.0	3138	14	AY406167 Homo sapi
2	5359	98.1	4775	6	CR859100 Pongo pyg
3	5116	93.6	3138	14	AY406168 Pan trogl
4	5054.5	92.5	3147	14	AY406169 Mus muscu

5	2916	53.4	3743	6	AK171713 Mus muscu
6	2896.5	53.0	3874	6	AK050860 Mus muscu
7	2318.5	42.4	3165	14	AY415515 Homo sapi
8	2286	41.8	3168	14	AY415517 Mus muscu
9	1920.5	35.2	3126	14	AY415516 Pan trogl
10	1648	30.2	968	4	BX364606 BX364606
11	1615	29.6	1000	4	BX354921 BX354921
12	1581.5	28.9	1354	10	DV784833 DV784833
13	1544	28.3	1026	2	BM551766 BM551766
14	1514	27.7	912	3	BQ881649 BQ881649
15	1496	27.4	904	3	BQ931960 BQ931960
16	1463.5	26.8	959	3	BUS39080 BUS39080
17	1448	26.5	1072	4	BX354920 BX354920
18	1445.5	26.5	916	3	BH846495 BH846495
19	1445.5	26.5	1043	3	BQ072025 BQ072025
20	1444	26.4	926	3	BQ706499 BQ706499
21	1442	26.4	857	4	BX50128 BX50128
22	1441.5	26.4	966	3	BUI46764 BUI46764
23	1418	26.0	889	3	BUI74449 BUI74449
24	1395	25.5	951	4	BX364605 BX364605
25	1383	25.3	806	10	DR763338 DR763338
26	1382.5	25.3	923	3	BQ950768 BQ950768
27	1380	25.3	2161	6	AK034903 AK034903
28	1373	25.1	800	9	CK781945 CK781945
29	1363.5	25.0	899	4	BX327170 BX327170
30	1357	24.8	896	3	BQ672632 BQ672632
31	1355	24.8	1122	7	BE746268 BE746268
32	1344	24.6	800	5	CD351400 CD351400
33	1341.5	24.6	881	3	BUI63030 BUI63030
34	1324	24.2	804	2	BG747792 BG747792
35	1321	24.2	774	2	B1760610 B1760610
36	1319.5	24.2	930	1	AL566821 AL566821
37	1319	24.1	861	2	BG764061 BG764061
38	1310	24.0	804	8	CN528131 CN528131
39	1306	23.9	784	2	BG763918 BG763918
40	1299	23.8	773	4	CB519600 CB519600
41	1294	23.7	2816	6	AK086699 AK086699
42	1289	23.6	952	3	BQ676994 BQ676994
43	1288.5	23.6	799	9	CX568240 CX568240
44	1287.5	23.6	865	7	BE260677 BE260677
45	1282	23.5	957	3	BUS20574 BUS20574

ALIGNMENTS

RESULT 1
AY406167
LOCUS
DEFINITION Homo sapiens FANP1 gene, VIRUTAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AY406167
VERSION AY406167.1 GI:39762141
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civeille,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Srinisky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civeille,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Srinisky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission

Db	1801	GAAGATTGAGCAACGACTTGGCCCTGTGGGGAAGGCCGCTCAATTGCCCAATTACAGAGATTAC	1860
Qy	621	GLNARGIIIEGIYAERVAIMEIEULEUYAHANIIIEGINIYMEIEUYHIELEUAIAIAHIE	640
Db	1861	CAAGAAATCGCGGAGATGATGCTGTAAGAACATTTCAGGCGCATGAGGACCTGGCGGCTTAC	1920
Qy	641	LEUTRPLYBHIESGILUALALEUGILUALALEUGIUAANGIYIIIEYBSESETHARGX	660
Db	1921	CTGTGGAAAGCACAGGAGGCGCTTGGAGGCGCTTGGAGATGGAATCMAGAGCTCCCGGCGG	1980
Qy	661	LEUGIUAENPHECYAARGASPHEGIULEUGIUNYVAIICYETRYLEUPROLEUANTHR	680
Db	1981	CTGGAGAACTTCTGACAGAGACTTTTGAAGCTTCGACAGAGAGTGTGTATTCTACCGCTCAACCC	2040
Qy	681	PHLEULEUARGPROLEUHIIEARGLEUWETHIERYTRYLYSGINVAIIEUGIUAIRLEUCYS	700
Db	2041	TTTCCTCTCGCGGCGCATCTCACCGGCTATGACATTAAGACAGAGGTCTTGGACGGCTGTGC	2100
Qy	701	LYBNHBNHIEPROPTROSETHIIEAIAERPHEARGARPCYAHGAIIAIALEUAIEGLUIE	720
Db	2101	AAACACCCACCGGCGGACGACGCGCATTTCAAGGACCTCCGACGCGCTTGTGGACAGATC	2160
Qy	721	THRGIMETVALAIEGLIUEHIEIEGLYTHMEIIEIYBMEIEGIIUAENPHEGINIYLEU	740
Db	2161	ACGGAGATGTGTGGCACACGCTCCACGGIACGATGATCAAGATGAGAGAAATTTCAAGAGCTG	2220
Qy	741	HIEGLIUEULYLYASAPLEUIIEGLYIIIEASPHANLEUVAIIPROGIYARGIUPHE	760
Db	2221	CACGAACTCAAGAAAGATTGATTGGATTGACAACTTGTGTGTTCGGGAAAGGAGATTC	2280
Qy	761	IIIEARGLEUGIYSEIRLEUSEIRLYALEUSEIRGIYLYSGIYLEUGINGIINARGMECPHEPHE	780
Db	2281	ATCCGCTCTGGCGCAGCTCAGACAGCTCTCGGGGAAAGGGGCTCCAGACGCGCATTTCTTC	2340
Qy	781	LEUPHEANAPVALIIEULEUTRYTHSERXRGIIYLEUTHRALASERAAHGINPHIYB	800
Db	2341	CTGTTCMAACGACGTCCTGTCTATACCGAGCGGGGGCTGACGGGCTCCAACTCACTTTAAA	2400
Qy	801	VAIIIEGLIYGINLEUPROLEUTRYCYIYMEETHRIIEGLIUGISERGIIUAERPGIUTRGIY	820
Db	2401	GTCCACGGGCGAGCTCCCGCTCTATGTGCATGACATTAGAGAGACGAAAGACGATGGGGG	2460
Qy	821	VALPROHIECYBLEUTHIRLEUARGIYGINARGIINSEIRIIEIEVAIAIAIASERSEIR	840
Db	2461	GTGCGCCCACTCCCTGACCCCTCCGGGGGCGACGGGAGTCCATCATCTGTGGCGGCAGTCT	2520
Qy	841	ARGSERGIUMETGIULYETRPIVALGIUAERPIIEGINMETHAIIIEAERPLEUAIAGIUNYS	860
Db	2521	CGGTCCGAGATGAGAAAGTGGGTGTGAGACATCCAGATGGCCATTGACCTTGGCGGAAG	2580
Qy	861	SEIRSEIRPROIALAPROGIUPHELEUIIASERSEIRPROAPAPANILYSEIRPROASP	880
Db	2581	AGCAGACAGCCCCCGCCCTGAGATTCTGTGGCCACGACCCCCCTTGACMAACAGTCCCCCTGAT	2640
Qy	881	GIUALATHRALAIAERPGIINGIUSEIRGIUAERAPLEUSERALASERARGITHSERIEU	900
Db	2641	GAAGCCACCGGCGGTGACACAGAGTCAGAGGATGACCTGACGCGCTCGCGCACATCCGCTG	2700
Qy	901	GIUARGIINAIAPROHIEARGIYASHTHMEIYAHIEVAIICYETRPHIEARGAENTHR	920
Db	2701	GAGGCGCCAGGCGCCGCGCACCGGCGGACACATGTGTGACGGTGTCTGGACCCGACACCC	2760
Qy	921	SEIRVALSERMETVALAERPHESERIIIEAIAVAIIGIUAENGIIUEUSERGIYASHTLEU	940
Db	2761	AGGCTCTTCATGTGTGACTTCAACGATCGCAATGAGAAATCACTGTGTCTGGAAACTCTCTG	2820
Qy	941	ARGIYSEIRPLEYASENSERANGIYTRPGIUNYALEUTRPIVALIIPHEITHASPHECYS	960
Db	2821	AGGAAATTCAAAAACAGCACCGGGGCGCAGAGCTGTGGGCGGTTCACAAATCTTGCC	2880
Qy	961	LEUPHERPEIYTRYBSEIRHIEGINAPRAPHNIEPROLEUIIASERIEUPROLEUIEUGIY	980
Db	2881	CTGTTCCTTCTACAAATCACACACAGACATATATCCCTTGGCACCTGCTGTGCTGGC	2940

Oy	981	TyrsertlewhrtlieProserGluSerGluAsnIleGlnIlysaPrrYrValPheIysIeu	1000
Db	2941	TACGGCTGCACCAATCCCTCTGAGTCCGAGAACATCCAGAAAGCATCGGTTCACACTG	3000
Oy	1001	HisPheIysSerHisValTYrTYrPheArgAlaGluSerGluTYrThrPheGluArgTrr	1020
Db	3001	CACTTCAAGTCCACGCTCTACTACTTACGAGGGCGGAAAGCAGATACAGTTCCAAAGGTGG	3060
Oy	1021	MetGluValIleArgSerAlaThrSerSerAlaSerArgProHisValIleSerHisIys	1040
Db	3061	ATGGAAGTGAATCCCGACATGGCACACGACTGTCCCTCGGACCCACCGTGTGATGCACAA	3120
Oy	1041	GluSerIeuValTYr 1045	
Db	3121	GAGTCTCTGTGTAT 3135	
RESULT 2			
LOCUS	CR859100	4775 bp	mRNA linear HTC 12-NOV-2004
DEFINITION	Pongo pygmaeus mRNA; CDNA DKFZp469C092 (from clone DKFZp469C092).		
ACCESSION	CR859100		
VERSION	CR859100.1	GI:55729110	
KEYWORDS	HTC.		
SOURCE	Pongo pygmaeus (orangutan)		
ORGANISM	Pongo pygmaeus		
REFERENCE	Bukaryotc, Metazoa, Chordata, Craniata, Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Pongo. 1 (bases 1 to 4775)		
AUTHORS	Bloecher, A., Boecher, M., Brandt, P., Mewes, H.W., Weill, B., Amlid, C., Oeanger, A., Fobo, G., Han, M. and Wiemann, S.		
CONSRM	The German cDNA Consortium		
TITLE	Direct Submission		
JOURNAL	Submitted (12-NOV-2004) MIPS, Ingolstaedter Landstr.1, D-85764 Neuherberg, GERMANY		
COMMENT	Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; Sequenced by GBR (National Research Centre for Biotechnology Ltd., Braunschweig/Germany) within the cDNA sequencing consortium of the German Genome Project. This clone (DKFZp469C092) is available at the RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany. Please contact RZPD for ordering: http://www.rzpd.de/cgi-bin/products/c1.cgi?cloneID=DKFZp469C092 Further information about the clone and the sequencing project is available at http://mips.gsf.de/projects/cdna/. Location/Qualifiers		
FEATURES			
source	1. .4775		
	/organism="Pongo pygmaeus"		
	/mol_type="mRNA"		
	/db_xref="taxon:9600"		
	/clone="DKFZp469C092"		
	/tissue_type="kidney"		
	/clone_id="469 (synonym: pkid1). Vector pSport1_sfi; host DH10B; sites SfiIa + SfiIb"		
	/dev_stage="adult"		
	/note="F8RM, RhogEF, and plectstrin domain protein 1(Homo sapiens) "		
gene	1. .4775		
	/gene="DKFZp469C092"		
CDS	100. .3237		
	/gene="DKFZp469C092"		
	/codon_start=1		
	/product="hypothetical protein"		
	/protein_id="CAH91292.1"		
	/db_xref="GI:55729111"		
	/translation="MGSEIORPPGSRIGAPENSGISTLERGQKPPPTSGKLSIKI OMIDPOBAFEVQPARAGKTLDAVCNHLIVGDFGECPPHKITWLDLKPVL KQIDRPYVYVKKVYKFPFPPDHOLOBELTRYFALOVQODLAQGRITCNDSSAALI SHVQSEIGDPDALDREHLAKKRYIPQDPALEDKIVYFHHNITGTPAESDPQLEI ARLEUQRLHAPAKOREGTILVAVANTILVFQGTAKINAMAKVRLSKRKF LITLRPANSAVQDTLEFLMASNDPFCSPKICIVENHAAPRLFEERPKVPVLSFG		

SSPFRSGRTOKOVLVDYVEGGKKVKVOFERKSKITHSIRLSASOPTLEYSEVLBOOS
ALUTFRGASBPGGSCCOQKEPRVSPGSEPHSESPVPRSPRANKNADGASAPTE
EEBVKRTQOSKROPPROPSTGLTSGPHLSHSVNSGCVAPRANSPDSTRO
ASPLISPLNDQACPTDDEDEGRKRPRPTDKAFIAEVSSTERTYTKDEVTSMF
QSAVSKEDAMPEAKSLIFNFEPLHKFHNFLKEIQRLAMWGRNAQIDRYQRG
DVMLKNTQGMKGLAVHLMKSEBALLENGLKSRRLNFORDELQKVCYLPANTFL
LRPLRHMVHKQVLERLCNHPPSHADFRCDRALAAGITTEWVAOLHGTMTIMENPQKL
HELKDLIGIDNLVVPGRREFIRGLSLSLGGKQOQMFLENVILLYTSRGLTASNO
FKVNGOLPLYMTGKTESPDEWGVPHCYLTIRBOROSITVAASRSEMEKRWDTOMAD
LAEKSSSLAEPFLASPPDNKSPDEBTRADDESDDISASTSLERAPRNGRTMAY
CMKRNVSVMVDFVAAVENQJSGNLRKFKNSNGMOKLWVFNPCLPFYKSHDNBP
LASLPLGLYSLTIPTESENNIHKDYVFKLHFKSHVYFRASEEYTFERMEVIRSATSS
ASRVHVSCHKESLVY"

ORIGIN

Alignment Scores:

Pred. No.:	0	Length:	4775
Score:	5359.00	Matches:	1025
Percent Similarity:	98.8%	Conservative:	7
Best Local Similarity:	98.1%	Mismatches:	13
Query Match:	98.1%	Indels:	0
DB:	6	Gaps:	0

US-09-555-342B-2 (1-1045) x CR859100 (1-4775)

Qy 1 MetGlyIuIleGluGlnArgProThrProGlySerArgLeuGluYAlaProGluAsnSer 20
Db 100 ATGGAGAAATAGACAGAGGCCAGCCCAAGATACAGACTGGGGGCCCAAGAAATTCG 159
Qy 21 GlyIleSerThrLeuGluArgGlyGluIleAspProProThrProSerGlyIleLeuVal 40
Db 160 GGATACAGTACTTGGAACGTGACAGAGCCGCCCAACCTTCAGAGAAATCGTGG 219
Qy 41 SerIleIleValIleGlnMetLeuAspAspThrGlnGluIleAlaPheGluValProGluArgAla 60
Db 220 TCCATCAAAATCCAGATGCTGATGACACCCAGAGGCGATTGAAATTCACAAAGAGCT 279
Qy 61 ProGlyIleValIleLeuAspAlaValCysAsnHisIleLeuAsnLeuValGluIleAspTyr 80
Db 280 CCGTGGAGAGTCTGCTGATGACAGTTCGAAACCACTTCACCTCGTGGAGGTGACTAT 339
Qy 81 PheGlyLeuGluPheProAspHisIleValIleThrValIlePheLeuAspLeuIleAspPro 100
Db 340 TTGGCCCTCGAGTCTCGATCACAAAGATCAACGGTGGTGGATCTCTTAAATCCC 399
Qy 101 IleValIleGlnIleArgArgProIleHisIleValIleValIlePheValIlePhePhe 120
Db 400 CTGTGAAACAGATCAGAGGSCCAAGATGTGTGTTAAGTTGTGTGTAATTCCTT 459
Qy 121 ProProAspHisIleThrGlnLeuGlnGluIleuThrArgTyrLeuPheAlaLeuGlnVal 140
Db 460 CCGGCTACCAACCACTCAAGAGAACTCAAGAGTACTCTGTTCGGCTCGACGGTG 519
Qy 141 LysGlnAspLeuAlaGlnGlyArgLeuThrCysAsnAspThrSerAlaAlaLeuIle 160
Db 520 AAGCAGAGACTTGGCTCAAGGCGAGTTCAGCTGTAACGACACAGCGACTCTTGAT 579
Qy 161 SerIleIleValIleGlnSerGluIleGlyAspPheAspGluAlaLeuAspArgGluHisIleu 180
Db 580 TCACACATTTGCAATCGAATGGGATTTTATGTAAGCTTGGACAGAGACACTTA 639
Qy 181 AlaIleAsnLeuIleGlyIleProGlnGlnAspAlaLeuGluAspIleValIleGluPheHis 200
Db 640 GCAAAATTAATATCATCTCAGCAAGACCACTAAGAGCAAAATCTGTGAATTCAC 699
Qy 201 HisAsnHisIleGlyIleThrProAlaGlnSerAspPheGlnLeuGluIleAlaArg 220
Db 700 CATTAACCACTTGGACAAACACAGCAAGATCAGATTTCCAGCTCTAGAGATTCGCCGT 759
Qy 221 ArgLeuGlnMetTyrGlyIleArgLeuHisProAlaIleAspAspArgGluGlyThrIleIle 240
Db 760 CGGCTAGAGATGTATGGAATCCGGCTGACCCGGCCAAAGACAGAGAAAGGACAGAAATC 819

Qy 241 AsnLeuAlaValAlaAsnThrGlyIleLeuValPheGlnGlyPheThrIleValAsnAla 260
Db 820 AATGGCCCGTTGGCAACAGGAAATCTAGTGTTCAGGGTTTCATTAAGATCAATGCC 879
Qy 261 PheAsnTrpAlaIleValIleArgIleLeuSerPheIleArgIleValArgPheLeuIleValLeu 280
Db 880 TTCAACTGGGCCCAAGGCGGAGAGCTGAGCTTCAAGAGAAAGCCCTTTCATCAAGCTC 939
Qy 281 ArgProAspAlaAsnSerAlaTyrGlnAspPheThrLeuGluPheLeuMetAlaSerArgAsp 300
Db 940 CGGCACATCCCAATAGTGGTACCAAGATTAACCTTGAATTCCTGATGGCCAGTCGGGAT 999
Qy 301 PheCysIleSerPheThrIleValIleCysValGluHisIleAlaPhePheArgLeuPheGlu 320
Db 1000 TTCTGCAAGTCCCTCTGGAATCTGTGTTGACATCATGCTCTTAAAGCTTTTGAA 1059
Qy 321 GluProIleAspProIleAspProValLeuPheSerArgIleSerSerPheArgPheSer 340
Db 1060 GAGCCCAAAACCAAGCCCAAGCCCGTCTCTTAAAGCCGGGGTTCATATTCGGTTCAGT 1119
Qy 341 GlyArgThrGlnIleGlnValIleAspTyrValIleGluGlyHisIleValIleGln 360
Db 1120 GGTGGACTCAGAAAGCAGTTCCTGACTATGTAAAGAGAGAGACATTAAGAGGTGACG 1179
Qy 361 PheGluArgIleHisSerIleHisSerIleArgSerLeuAlaSerGlnProThrGlu 380
Db 1180 TTGGAAGAGAGACACAGCAAGATTCATTCATCCGAGCCTTGCTTCAAGCCTTAAGAA 1239
Qy 381 LeuAsnSerGluValIleGlnGlnSerGlnGlnSerThrSerLeuThrPheGlyGluGly 400
Db 1240 CTGTATTCGGAAGTACTGAGAGTCTCAGCAGAGCCCAAGCCTTCAATTTGGAGAAAGT 1299
Qy 401 AlaGluSerProGlyIleGlnSerCysArgArgGlyIleGluIleAspIleValIleGly 420
Db 1300 GCCAATCTTCAAGGGGCCAGAGTCTCCACAGAAAGAAAGAACCAAGGTTTCCCGCGG 1359
Qy 421 GluProGlySerHisProSerProAlaProArgArgSerProAlaGlyAsnIleGlnAla 440
Db 1360 GAGCCGGGGTGGACCCGAGCCCTGTGTCGAGAGAAAGCCCGGGTTAAACAGAGCGG 1419
Qy 441 AspGlyAlaAlaSerAlaProThrGluGluGluGluValIleValAspArgThrGln 460
Db 1420 GACGAGACCCCTCGGGCCCAAGAGAGAGAGAGGTGTGTAAGATGAGACCAAG 1479
Qy 461 GlnSerIleProGlnProProGlnProSerThrGlySerLeuThrArgIleSerProHisIleu 480
Db 1480 CAGAGTAACTCAGCCCGGACGACCAAGCAGAGCTCCCTGACCGGACGCCCTCACCTT 1539
Qy 481 SerGluLeuSerValaAsnSerGlnGlyIleAlaProAlaAsnValIleThrLeuSerPro 500
Db 1540 TCTGAGCTGTCTGGAATCGCAGGGGGAGTGGCCCTCGCAACGTGACTGTCTCC 1599
Qy 501 AsnLeuSerProAspThrIleGlnAlaSerProLeuIleSerProLeuLeuAsnAspGln 520
Db 1600 AACCTGACCCCAACACAGAGGCTCTCCCTGTATCAGCCCGCTGTGAAGACACAG 1659
Qy 521 AlaCysProArgThrAspAspGluAspGluGlyArgArgIleAspPheProThrAspIle 540
Db 1660 GCTCGCCCGGACGAGCAGTGAAGATGAGGGCCGAGAGAAAGATTCACCACTGATTA 1719
Qy 541 AlaTyrPheIleAlaIleGluValIleSerThrThrGluArgThrTyrIleuIleAspLeuGlu 560
Db 1720 GCGTACTTCAATAGTAAGAAAGTGTACACCAAGAGCAATATTCGAAGATCTGAA 1779
Qy 561 ValIleThrSerTrpPheGlnSerThrValIleSerIleGluAspAlaMetProGluAlaLeu 580
Db 1780 GTTATCACTTCGGTTCAGAGCGCAGTAGCAAAAGAGACGCAATCCAGAAAGCACTG 1839
Qy 581 LysSerLeuIlePheProAsnPheGluProLeuHisIleYspPheHisThrAsnPheLeuIle 600
Db 1840 AAAGTCTCATATTCGCCAATTTTGAACCTTTTGCAAAATTTCTATCAATTTTCTCAG 1899
Qy 601 GluIleGlnIleArgLeuAlaLeuTrpGluGlyArgSerAsnAlaGlnIleArgAspTyr 620


```

Db      1900 GAAATTGACGACGCTTGCCTGTGGAGGCGCCTCAAAATGCCAAATCAGATTTAC 1959
Qy      621 Glnatggttgglyabpvalmetleulyabnllleglmglymetlyhlsleualalahlis 640
Db      1960 CAAGAATTCGGGAGTGTATGCTGAAACAATTCAGGGGATGAAACACCTGGCGGTTTAC 2019
Qy      641 Leutrlpshhsersglualaleuglualaleuglualaleuglualaleuglualaleugl 660
Db      2020 CTGTGGAAACAGCAGGAGCTTGGAGGCGCTGGAGAAATGGATCAAGAGCTCCGCGCGG 2079
Qy      661 LeugluabepheCyargAapPhegluleuglnlyvalCystryleuProleuanthr 680
Db      2080 CTGGAGAACTTCTGACAGACTTTAGAGCTGCAGAAAGTGTGTACTTACCTGCAACACC 2139
Qy      681 PheuleuargProleuHlsargleuethslyrlyrlyglvalleuglnargleuCy 700
Db      2140 TTCCCTCTGGCGGCACTGCACCGGCAATGCACCTCAAGCAGGTCTGTGAGCGGCTGTGC 2199
Qy      701 LysHshlsapropserHlsalAapPheargAapCyargAlalaleualaleuallile 720
Db      2200 AAACACCAACCGGCGAGCCAGCCGACCTTCAGGGAGAGGGGCTTCACAGCGCATGTTCTC 2259
Qy      721 ThrglumetvalaleglleuHlslyrThrmecHlslysmecGlualanPheglulyleu 740
Db      2260 ACGGAGATGTGGCAGCAGCTCCAGGTAAGATCAAGATGAGAAATTTCCAGAAAGCTG 2319
Qy      741 HlsleuleulylyeAapPleuHlslyrlyleapAanleuValproglAarggluPhe 760
Db      2320 CACGAACTCAAGAAAGATTTGATTGATTCATTCATTCGTTGCTCCGGAAGGAGATTTC 2379
Qy      761 HlsargleuglyserleuSerlyrleuSerlyrlyrlyrlyrlyrlyrlyrlyrlyrlyr 780
Db      2380 ATCCGTCTGGGAGGCTCAGCAAGCTCTCGGGAGAGGGGCTTCACAGCGCATGTTCTC 2439
Qy      781 LeupheAanAapvalleuLeuTyrrThserAarglyleuThnAlaserAasnlglnPhe 800
Db      2440 CTGTTCACAGAGCTCTCTGTACACAAAGCCGGGGCTGCACGCTCCATCAGTTTAA 2499
Qy      801 ValHlslyglleuProleuTyrrlyrlyrlyrlyrlyrlyrlyrlyrlyrlyrlyrly 820
Db      2500 GTCCACGGCCAGCTCCGCTCTCAACGAGATGAGATTAAGAGAGGAGAAACAGAGGGGG 2559
Qy      821 ValproHlsCybleuThleuarglyrlyrlyrlyrlyrlyrlyrlyrlyrlyrlyrlyr 840
Db      2560 GTGCCCCAAGCTGAGCCCTCCGGGGCCAGCGGCACTCATCATGTTGCGCGCCAGTTCT 2619
Qy      841 Argserglumetglulyrlyrlyrlyrlyrlyrlyrlyrlyrlyrlyrlyrlyrlyrlyr 860
Db      2620 CGGTCTGAGATGAGAAAGTGGGTGAGAGATCCAGATGGCCATTGACCTGGCAGAGAG 2679
Qy      861 SerSerSerProAlaProgluPheleuHlsaserSerProAapAanlyaserProAap 880
Db      2680 AACAGCAACCTCTGCGCTGAGATTCTGTGGCCAGCAACCCCTGACAAACAAGTCCCTGAT 2739
Qy      881 GlualatrrAlaAlaAapgluHlsSerGlualapAapleuSerAlaserAargThserleu 900
Db      2740 GAAGCACCGGCGCTGACAGAGATCAAGATGAGCTGAGGCGCTTCGCGCACATCGCTG 2799
Qy      901 GlualrgluAlaProHlsAarglyrlyrlyrlyrlyrlyrlyrlyrlyrlyrlyrlyr 920
Db      2800 GAGCGCCAGGCGCCGCAACCTGGCAACAAATGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2859
Qy      921 SerValSerMetValAapPheSerHlsAlaValGlualanGlnleuSerGlyAanleu 940
Db      2860 AGCGTCTCATGTGTGAGCTTCAGCTGTGAGTGTGAGAAATCATGTTGTCTGGGAACCTGCTG 2919
Qy      941 ArglyrPheleuAanSerAasnlyrlyrlyrlyrlyrlyrlyrlyrlyrlyrlyrlyrlyr 960
Db      2920 AGGAATTCGAAACAGCAACGGGTGGCAAGAGCTGTGGGTGTGTGTGTGTGTGTGTGTGTGT 2979
Qy      961 LeuphePheTyrrlyrlyrlyrlyrlyrlyrlyrlyrlyrlyrlyrlyrlyrlyrlyrlyr 980

```

```

Db      2980 CTGTCTCTTAATGATCACACGAGCATCATCCCTCGCAGCTGCTGTGCTCGGC 3039
Qy      981 TyrrSerleuThrlleProSerGluserGluserlnleuglnlyrlyrlyrlyrlyrlyrlyr 1000
Db      3040 TACTGCTCACATCCCACTGAGTCCAGAAACATCCAAAGACTTACGTTCAGAGCTG 3099
Qy      1001 HlsPheylsSerHlsvalTyrrlyrlyrlyrlyrlyrlyrlyrlyrlyrlyrlyrlyrlyr 1020
Db      3100 CACTTCAAGTCCCACTTACTTACTTACGCGGCGGAAACAGATACACTTGAAGAGTGG 3159
Qy      1021 MetgluValHlsargSerAlaThrserserAlaserAapProHlsvalleuSerHlslys 1040
Db      3160 ATGGAAGTATCCGAGCGGCGCAGCTGTGCTGTGAGTGTGAGTGTGAGTGTGAGTGTG 3219
Qy      1041 GluserleuValTyrr 1045
Db      3220 GAGTCTCTGTGTAT 3234

RESULT 3
AY406168
LOCUS      Pan troglodytes FAR1 gene, VIRTUAL TRANSCRIPT, partial sequence,
DEFINITION genomic survey sequence.
ACCESSION AY406168
VERSION    AY406168.1 GI:39762142
KEYWORDS   GSS.
SOURCE      Pan troglodytes (chimpanzee)
ORGANISM   Pan troglodytes
REFERENCE   1 (bases 1 to 3138)
AUTHORS    Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,
            Todd,M.A., Tenenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
            Inferring nonneutral evolution from human-chimp-mouse orthologous
            gene trios
JOURNAL    Science 302 (5652), 1960-1963 (2003)
PUBMED     14671302
REFERENCE   2 (bases 1 to 3138)
AUTHORS    Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,
            Todd,M.A., Tenenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
            Direct Submission
JOURNAL    Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT     This sequence was made by sequencing genomic exons and ordering
            them based on alignment.
FEATURES
            source
            1..3138
                /organism="Pan troglodytes"
                /mol_type="genomic DNA"
                /db_xref="taxon:9598"
                <1..3138
                /gene="FAR1"
                /locus_tag="HCM2467"
ORIGIN
Alignment Scores:
Pred. No.:      0          Length:      3138
Score:          5116.00    Matches:      989
Percent Similarity: 94.8%   Conservative: 2
Best Local Similarity: 94.6% Mismatches: 54
Query Match:    93.6%      Indels:      0
DB:             14        Gaps:      0
US-09-555-342B-2 (1-1045) x AY406168 (1-3138)
Qy      1 MetgluValHlsargSerAlaThrserserAlaserAapProHlsvalleuSerHlslys 20
Db      1 ATGGAGAAATTAAGACAGAGCCGACCCCAAGATCGGACTGGGGGCCCCGGAATAATTCG 60

```

QY 21 GlyIleSerThrLeuGluArgGlyGlnIysProProThrProSerGlyIysLeuVal 40
Db 61 GGAGTCACTGACCTTGGAAAGCGTGCACAGAGCCGCCCAACACTTCAGGAAATCTGTG 120
QY 41 SerIleIysIleGlnMetLeuAspAspThrGlnGluAlaPheGluValProGlnArgAla 60
Db 121 TCCATCAAAATCCAGATGCTGTGATGACACCCAGAGGCGATTGTGAAGTTCCANNNNNNNN 180
QY 61 ProGlyIysValLeuLeuAspAlaValCysAsnHisIleuLeuValGluGlyAspTyr 80
Db 181 NNN 240
QY 81 PheGlyLeuGluPheProAspHisIleIysIleThrValTrpLeuAspLeuLeuIysPro 100
Db 241 NNN 300
QY 101 IleValIysGlnIleArgArgProIysHisValValIysPheValValIysPhePhe 120
Db 301 ATTGTGAACAGATTAGAGAGCCAAAGCATGTGTGTGAAGTTGTGTGTAATTTCTTT 360
QY 121 ProProAspHisThrGlnIleuGlnGluIleuThrArgTyrIleuPheAlaLeuGlnVal 140
Db 361 CCGCTGACCAACACAACTCCAAAGAACTCAAGAGTACCTGTTCGGCTCGAGGTG 420
QY 141 LysGlnAspLeuAlaGlnGlyArgLeuThrCysAsnAspThrSerAlaAlaLeuIle 160
Db 421 AAGCAGAGCTTGGCTCAAGCAGGTTGACGTGTAATGACACACAGCAGCTCTTGATT 480
QY 161 SerHisIleValGlnSerGluIleGlyAspPheAspGluAlaLeuAspArgGlnHisIleu 180
Db 481 TCACACATTTGCAATCTGAGATTGGGGATTGTGATGAGCCTTGAGCAGAGACACTTA 540
QY 181 AlaIysAsnIysTyrIleProGlnGlnAspAlaLeuGluAspIysIleValGluPheHis 200
Db 541 GCAAAATTAATATCATTACTCAGACAAAGCCACTAGAGCAAAATCGTGGAAATTCAC 600
QY 201 HisAsnHisIleGlyGlnThrProAlaGlnSerAspPheGlnIleuLeuGluIleAlaArg 220
Db 601 CATATCAACATTTGCAAAACACAGCAGAAATCAATTTCCAGCTTCAGAGATTGCCGT 660
QY 221 ArgLeuGlnMetTyrGlyIleArgLeuHisProAlaIysAspArgGluGlyThrLysIle 240
Db 661 CGGTAGAGATGATGAGATCCGGTTCACCCGGCCAAAGCAGAGGAGGACCAAAATC 720
QY 241 AsnLeuAlaValAlaAsnThrGlyIleIleuValPheGlnGlyPheThrLysIleAsnAla 260
Db 721 AATCTGGCCGTTGCCAACAAGGAAATCTAGTGTTCAGGGTTTCACCTAAGATCAATGCC 780
QY 261 PheAsnTrpAlaIysValArgIysLeuSerPheIysArgIysArgPheLeuIleIysLeu 280
Db 781 TTCAACTGGGCGCAAGCGTGGGAGCTGAGCTTCAAGAGGAGGCCCTTTCATCAAGCTC 840
QY 281 ArgProAspAlaAsnSerAlaTyrGlnAspThrIleuGluPheLeuMetAlaSerArgAsp 300
Db 841 CGGCAATGCGCAATAGTGCCTACAGGATTCCTTGAATTCCTGAATGGCCAGTCCGGAT 900
QY 301 PheCysIysSerPheTrpIysIleCysValGlnHisIleAlaPheArgLeuPheGlu 320
Db 901 TTCTGCAAGTCTCTTGAAATCTGTGTGAAACATACAGCTCTTTAAGCTTTTGA 960
QY 321 GluProIysProIysProIysProValLeuPheSerArgIysSerPheArgPheSer 340
Db 961 GAGCCCAACCAAGCCCAAGCCCGTCTTTAGCCGGGGTCAATCATTTCCGTTCACT 1020
QY 341 GlyArgThrGlnIysGlnValLeuAspTyrValIysGluGlyIysIleIysValGln 360
Db 1021 GGTTCGAGCTCAGAGAGGAGTTCTGACTATGTAAAGAGAGGACATTAAGAGGTCCAG 1080
QY 361 PheGluArgIysHisSerLysIleHisSerIleArgSerLeuAlaSerGlnProThrGlu 380
Db 1081 TTTCAAAGGAAGCAGCAGAGATTCAATTCGAGAGCTTGCTTCAAGCTTACAGAA 1140

QY 381 LeuAsnSerGluValLeuGlnIysSerGlnIysSerThrSerLeuThrPheGlyGluGly 400
Db 1141 CTGAATTCGGAAGTGTGAGAGAGTCTCAGACAGGCCCAAGCTTACATTTGAGAAAGT 1200
QY 401 AlaGlnSerProGlyGlyGlnIysSerCysArgArgGlyIysGluProIysValSerAlaGly 420
Db 1201 GCCGAAATCTCAGAGGGGCCAGAGCTGCGGCAAGAAAGAAACGGAAGTTTCCGGCGG 1260
QY 421 GluProGlySerHisProSerProAlaProArgArgSerProAlaIysAsnIysGlnAla 440
Db 1261 GAGCCGGGGTTCGACCCGAGCCCTGCGCCGAGAGAAAGCCCGGGGTAAACAGCAGCG 1320
QY 441 AspGlyAlaAlaSerAlaProThrGlnGluGlnGluValIysValIysAspArgThrGln 460
Db 1321 GACGAGCCGCTCGGGGCCCAAGAGAGAGAGAGAGTCCGTTAAGATAGAGCAG 1380
QY 461 GlnSerIysProGlnProProGlnProSerThrGlySerLeuThrGlySerProHisIleu 480
Db 1381 CAGAGTAACTCAGAGCCCGCAGCCAGCAGCAGAGCTCCCTGACTGGCAGTCTCACCTT 1440
QY 481 SerGluIysSerValAsnSerGlnGlyValAlaProAlaAsnValThrLeuSerPro 500
Db 1441 TCCGAGCTGTCTGAACTMGACAGGGGGAGTGGCCCTGCAACGTGACTTCTCC 1500
QY 501 AsnLeuSerProAspThrIysGlnAlaSerProLeuIleSerProIleuAsnAspGln 520
Db 1501 AACCTGAGCCCGACACCAAGCAGAGCTCTCCCTGATCAGCCCGCTGAAATGACAG 1560
QY 521 AlaCysProArgThrAspArgGluAspGluGlyArgArgIysArgPheProThrAspLys 540
Db 1561 GCTGCCCCCGAGCGAGCGATGAGAGTGAAGGCGCGAGGAAAGATTTCCAACTGAATAA 1620
QY 541 AlaTyrPheIleAlaIysGluValSerThrThrGlnArgThrTyrLeuIysAspLeuGlu 560
Db 1621 CGGTACTTCACTAGCTAAGAGANNNTCCNNNACCAAGAGAACATATCGAAGATCTGAA 1680
QY 561 ValIleThrSerTrpPheGlnSerThrValSerIysGluAspAlaMetProGluAlaLeu 580
Db 1681 GTTATCACTTCGTGTGTTTCAAGACACAGTGAAGAAAGAGACCCATCGCGGAAGCACTG 1740
QY 581 LysSerIleuIlePheProAspPheGluProLeuHisIysPheHisThrAsnPheLeuLys 600
Db 1741 AAAAGTCAATATCCGAATTTGAACCTTGACAAATTTCAATTAATTTCTCAG 1800
QY 601 GluIleGlnIysArgLeuAlaLeuTrpGluGlyArgSerAsnAlaGlnIleArgAspTyr 620
Db 1801 GAATATGAGCAAGACTTGCCTGTGGAGAGCGCGCTCAATATGCCAAATCAAGATTC 1860
QY 621 GlnArgIleGlyAspValMetLeuIysAsnIleGlnGlyMetLysHisIleuAlaHis 640
Db 1861 CAAGAATCGCGCATGTCACTGAAGAACATTCAGGGCATGAAGCACCTGCGCGCTCAC 1920
QY 641 LeuTrpLysHisSerGluAlaLeuGluAlaLeuGluAsnGlyIleLysSerSerArgArg 660
Db 1921 CTGTGAAGCAGCAGAGGCTGTGAGGCTTGAAGTGAAGATCAAGAACTCCCGCGG 1980
QY 661 LeuGluAsnPheCysArgAspPheGluLeuGlnIysValCysTyrIleuProLeuAsnThr 680
Db 1981 CTGAGAACTTCTGCAAGACTTGAAGCTGCAAGAGTGTATTAACCGCTCAACACC 2040
QY 681 PheLeuLeuArgProIleuHisArgLeuMetHisTyrLysGlnValLeuGluArgLysCys 700
Db 2041 TTCTCTCTCGGCGACGTGCACCGGCTCATGCAANNACAAAGCAGTCCGTGAGCGGNNGTGC 2100
QY 701 LysHisIleProProSerHisIleAspPheArgAspCysArgAlaAlaLeuAlaGluIle 720
Db 2101 AAACACACCCCGCGGANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNCCGTTTGGCAGAAATC 2160
QY 721 ThrGlnMetValAlaGlnIleuHisGlyThrMetIleLysMetGluAsnPheGlnIysLeu 740
Db 2161 ACGGAGATGATGACAGCTCCAGCGGTACGATATCAAGATGAGAAATTTCCAGAAAGCTG 2220
QY 741 HisGluLeuLysIysAspLeuIleGlyIleAspAsnLeuValValProGlyArgGluPhe 760

```

Db      2221 CACGAACCTCAAGAAAGATTGATTGATGATGATGATCTTGCGTCCGGGAAAGGAGATTC 2280
Qy      761  ILeArgLeuGlySerLeuSerLeuSerGlyLeuGlyLeuGlnGlnArgMetPhePhe 780
Db      2281 ATCCGCTCTGGGACAGCTCAGCAAGCTCTCGGGGAAAGGGGCTCCAGCACGCGATGTTCTTC 2340
Qy      781  LeuPheAaAaPValLeuLeuTyrThrSerArgGlyLeuThrAlaSerPheGlnPheLeu 800
Db      2341 CTGTTCAACGAGAGTCTCTATACACGAGCCGGGGGCTGACGGCTCCAAATCAGTTTAA 2400
Qy      801  ValHISGlyGlnLeuProLeuTyrGlyMetThrIleGlnGlnSerGlyAaPgluTrrpGly 820
Db      2401 GTCCACGGGAGAGCTCCGCTCTATGCGATGACGATGAGACAGGAAAGAGAGGGGG 2460
Qy      821  ValProHISGlyLeuThrLeuArgGlyGlnArgGlnSerIleIleValAlaAlaSerSer 840
Db      2461 GTGCCCTCACTGCTGACCTCCGGGGCCAGCGGCGAGTCATCATCTGCGCCGCAAGTTCT 2520
Qy      841  ArgSerGluMetGluValTrrpValGluAaPgluIleGlnMetAlaIleAaPheAaGluLeu 860
Db      2521 CGGTCCGAGATGAGAAAGTGGGTGAGGACATCCAGATGGCCATTCGACCTGGCGAGAG 2580
Qy      861  SerSerSerProAlaProGluPheLeuAlaSerSerProProAaPheAaPheProAaP 880
Db      2581 AGCAGCAGCCCGCCCGCTGAGATTCCTGGCCAGCAGCCCGCTGACACAAAGTCCCTGAT 2640
Qy      881  GluAlaThrAlaAlaAaPgluGlnSerGluAaPheLeuSerAlaSerArgThrSerLeu 900
Db      2641 GAAGCCACCGCGGCTGACCAAGAGTCAAGAGATGACCTGAGCGCTCCGCGACGTCGCTG 2700
Qy      901  GluArgGlnAlaProHISArgGlyAaPheThrMetValHISValCysTrrpHISArgAaPhe 920
Db      2701 GAGGCGCAGGCGCCGCGACCGGCAACAAATGTCACGTCGTGCTGGCACCAGCAACCC 2760
Qy      921  SerValSerMetValAaPheSerIleAlaValGluAaGlnLeuSerGlyAaPheLeu 940
Db      2761 AGCGCTCTCAGTGGTGAATTCAGCATCCGAGTGGAGATCAGTTCTCTGGAACCTGCTG 2820
Qy      941  ArgGlyPheLeuAaPheSerAaPheGlyTrrpGlnLeuLeuTrrpValPheThrAaPheCys 960
Db      2821 AGGAAATTCAAAAACAGACACGCGGTGGCAAGAGCTGGTGGTGTCTTCAAAACTTCTGC 2880
Qy      961  LeuPhePheTrrpLeuSerHISGlnAaPheAaPheHISProLeuAlaSerLeuProLeuGly 980
Db      2881 CTGTTCTTCAAGTCAACAGTCAACAGGACCAATATATCCCTTGGCAGCTCTCTGCTCGGC 2940
Qy      981  TyrSerLeuThrIleProSerGluSerGluAaPheGlnLeuAaPheTrrpValPheLeu 1000
Db      2941 TACTCGCTCACATCCCTCCGAGTCCGAAACATCCGAAAGACTACGTGTTCAAGCTG 3000
Qy      1001  HISPheLeuSerHISValTrrpTrrpAaPheAaPheGlnSerGlyTrrpPheGlnArgTrrp 1020
Db      3001 CACTTCAAGTCCCAAGTCTACTACTTCAAGGGGCGGAAACGAGTACAGTTGGAAGGTGG 3060
Qy      1021  MetGluValIleAaPheSerIleThrSerSerAlaSerAaPheProHISValLeuSerHIS 1040
Db      3061 ATGAAAGATCCGAGTGCACACGACGACGACGACGACGACGACGACGACGACGACGAC 3120
Qy      1041  GluSerLeuValTrrp 1045
Db      3121 GAGTCTCTCGTGTAT 3135

```

```

REFERENCE
AUTHORS
1 (bases 1 to 3147)
Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarival, A.,
Todd, M.A., Tenenbaum, D.M., Clivello, D.R., Lu, F., Murphy, B.,
Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Shinsky, J.J.,
Adams, M.D. and Cargill, M.
TITLE
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene clones
JOURNAL
Science 302 (5652), 1960-1963 (2003)
PUBMED
14671302
REFERENCE
AUTHORS
2 (bases 1 to 3147)
Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarival, A.,
Todd, M.A., Tenenbaum, D.M., Clivello, D.R., Lu, F., Murphy, B.,
Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Shinsky, J.J.,
Adams, M.D. and Cargill, M.
TITLE
Direct Submision
JOURNAL
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT
This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source
1..3147
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
<1..>3147
/gene="FARP1"
/locus_tag="HMC2467"
ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 3147
Score: 5054.50 Matches: 962
Percent Similarity: 94.8% Conservative: 32
Best Local Similarity: 91.8% Mismatches: 51
Query Match: 92.5% Indels: 3
DB: 14 Gaps: 2
US-09-555-342B-2 (1-1045) x AY406169 (1-3147)
Qy      1  MetGlyIuIleGlnGlnArgProThrProGlySerArgLeuGlyAlaProGluAaPhe 20
Db      1  ATGGAGAAATTAAGACAGAACCCACCCACATCACTCTGGGGGCCCGAGAAACTCG 60
Qy      21  GlyIleSerThrLeuGluArgGlyGlnAaPheProProThrProSerGlyLeuVal 40
Db      61  GGATCAGACACTTGGAAACGAGACAGAAACACCGCCACACTTCAGAAAACCTTATG 120
Qy      41  SerIleLeuIleGlnMetLeuAaPheThrGlnGlnAlaPheGlnValProGlnArgAla 60
Db      121  ACTGTCAAAATTCAGATGCTGAGACACCCAGAGAGGCTTTGAAGTTCACACAGACG 180
Qy      61  ProGlyLeuValLeuLeuAaPheAlaValCysAaPheHISLeuAaPheValGluGlyAaPhe 80
Db      181  CAGGAGAGGTTCTGTTTGAACGCGGTGCAACCACTCAACTTGTGAGGGGCACTAC 240
Qy      81  PheGlyLeuGluPheProAaPheLeuValLeuValIleThrValTrrpLeuAaPheLeuAaPhe 100
Db      241  TTCGGCTGAGATTCCCTGACCAACAGAAAGATCGGGGTGGCTGAGATCTCTGAAAGCC 300
Qy      101  IleValIleGlnIleAaPheProLeuHISValValValIlePheValValIlePhePhe 120
Db      301  ATGTGAAGCAGATTTCGAAGACGAGAGACGTCGTGTAAATTTGTGAAATTTCTTC 360
Qy      121  ProProAaPheHISThrGlnLeuGlnGlnGluLeuThrArgTrrpLeuPheAlaLeuGlnVal 140
Db      361  CTTCCGACCAACACAGACTCCAGAAAGACTTACAGAGTACCTGTTGCTCTGCAAGTG 420
Qy      141  LysGlnAaPheLeuAlaGlnGlyArgLeuThrCysAaPheThrSerAlaIleLeuLeuIle 160
Db      421  AAGCAGAGCTTGCGCTCAAGGCGAGGCTGACGTAAACGACACCAAGTGGCGCTCTTGATC 480

```

QY 161 SerHisIleValGlnSerGluIleGIYAAspPheAspGluAlaLeuAspArgGluHisIleu 180
DB 481 TCACACATTGACGAGTGGAGATTGGGATTCATGAGCTTTGGACAGAGAACACTTA 540
QY 181 AlaIysAsnIleYrIleProGlnGlnAspAlaLeuGluAspLysIleValGluPheHis 200
DB 541 GCAGAAAATATAGTACCTCAACAGGATCGCTGGAGACAGAACTCATGAGTTCCAC 600
QY 201 HisAsnHisIleGlyGlnThrProAlaGlnSerAspPheGlnLeuGlnIleAlaArg 220
DB 601 CACAGCCAGCTGGACAAACCCCGCAGAGTCAATTTCCAGCTTCAGAGGTGGCCCGT 660
QY 221 ArgLeuGlnMetYrGlyIleArgLeuHisProAlaLysAspArgGluGlyThrIleVal 240
DB 661 CGGCTTAGATGTCGCGAATCCGGTTACCCCTCCAGACAGACAGAGAGCCACCAAGATC 720
QY 241 AsnLeuAlaValAlaAsnThrGlyIleLeuValPheGlnGlyPheThrLysIleAsnAla 260
DB 721 AACCTGGCTGTCGCAACACTGGAATTCATGTTTCAAGGCTTCACGAGATCATATGCT 780
QY 261 PheAsnThrAlaLysValArgLysLeuSerPheLysArgLysArgPheLeuIleLysLeu 280
DB 781 TTTTAACTGGGCGCAAGTGGCGAAGCTGAGCTTCAAGCGGAAGCGCTTCTTATCAAGCTC 840
QY 281 ArgProAspAlaAsnSerAlaYrGlnAspThrLeuGluPheLeuMetAlaSerArgAsp 300
DB 841 AGACCTACCGAATAGCTACTACCAAGACACTTTGAGATTCCTGATGGCGGCGGAGAC 900
QY 301 PheCysLysSerPheThrLysIleCysValGlnHisIleAlaPheArgLeuPheGlu 320
DB 901 TTTCGCAAGCTCTTCTGAAAAATATGTGTGAGACCACTGCTTCAAGCTCTTGAA 960
QY 321 GluProLysProLysProLysProValLeuPheSerArgLysSerSerPheArgPheSer 340
DB 961 GAACCTAAGCCAAAGCCAAAGCCGTGTTCTTTCAGCCGAGGGTGCTTCCGCTTAACT 1020
QY 341 GluYrGThrGlnLysGlnValLeuAspYrValLysGlnGlyGlyHisLysValGln 360
DB 1021 GGTGCGACTCGAAGCAAGTTCTTGATTTATTTAAAGCGGCGCCACAAAGAGTCTCAG 1080
QY 361 PheGluArgLysHisSerLysIleHisSerIleArgSerLeuAlaSerGlnProThrGlu 380
DB 1081 TTTGAAAGAAAGCACACCAAGATCCACTCCACAGAGAGCTGCTCAAGCCACAGCA 1140
QY 381 LeuAsnSerGluValIleGlnGlnSerGlnGlnSerThrSerLeuThrPheGlnGlnGly 400
DB 1141 CCGAATTCAGAAAGTGCAGAAACAGTCTCCACAGAGCCCAAGCTTACGTTGAGAAAGT 1200
QY 401 AlaGlnSerProGlnGlyGlnSerCysArgArgGlyLysGluProLysValSerAlaGly 420
DB 1201 ACCGAGCTCCGGGTGGCCAGAGCTGCCACAGAACGAAAGAAACAAAGCTGTAACCTTG 1260
QY 421 GluProGlySerHisProSerProAlaProArgArgSerProAlaGlyAsnLysGlnAla 440
DB 1261 GAGTGGGAGACTCCCAAGAGCCCGCCCTCCCTCAAGGAGAGCCCTCAAGAGCTGCA 1320
QY 441 AspGlyValAlaAsnAlaProThrGlnGlnGlnGlnValVal-----LysAspArg 458
DB 1321 GATGGAACCAACCGTGGCGCTCGAGAGAGAAAGAGAGAGGGGTGGCAAGAGAGCGC 1380
QY 459 ThrGlnSerLysProGlnProProGlnProSerThrGlySerLeuThrGlySerPro 478
DB 1381 ATCCGGGCAAGTAACTCCAGCCCGCCGACAGCAAGCAAGGCTCCGATCGATGAGCCCT 1440
QY 479 HisLeuSerGluLeuSerValAsnSerGlnGlyValValAlaProAlaAsnValThrLeu 498
DB 1441 CACCTTTCAAGGCTGCTCATCACTCAAGGAGAGAGCCCGCCAGCAAGTGAACCTTG 1500
QY 499 SerProAsnLeuSerProAspThrLysGlnAlaSerProLeuIleSerProLeuLeuAsn 518
DB 1501 TCTCCCAACTGAGTCTTGACAAACAGAGGCTCTCCCTCGATCAAGCCGCTCTGTAAC 1560

QY 519 AspGlnAlaCysProArgThrAspAspGluAspGlnGlyValArgArgLysArgPheProThr 538
DB 1561 GACCAACCGTCCCAAGAGCCGACGACGAGAGAGGCGCGGCAAGAGATTCACCAACC 1620
QY 539 AspIleValYrPheIleAlaLysGlnValSerThrThrGluArgThrYrLeuLysAsp 558
DB 1621 GACAGCGCTGATCATATGCTTAAGAGAGTCTCCACCAAGAGCGCACGTCAAGAGAC 1680
QY 559 LeuGlnValIleThrSerThrPheGlnSerThrValSerLysGlnAspAlaMetProGlu 578
DB 1681 CTGAGGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1740
QY 579 AlaLeuLysSerLeuIlePheProAsnPheGluProLeuHisLysPheHisThrAspPhe 598
DB 1741 GCCTTGAAGAGTCTCATATTCCAAAATTTGAACCTTTGCAACAGTTCATCAAAATTTT 1800
QY 599 LeuLysGlnIleGlnGlnArgLeuAlaLeuTyrGlnGlyValArgSerAsnAlaGlnIleArg 618
DB 1801 CTGAGGAAATTTGAGCAACAGACTTGCCCTGGGAGAGCGCTCAAAATGCCACATGAGA 1860
QY 619 ---AspYrGlnArgIleGIYAAspValMetLeuLysAsnIleGlnGlyMetLysHisIleu 637
DB 1861 GAGATTTACCAAGAAATCGGGAGCTCATGCTGAAAACATTCAGGGCATGAGACACTTG 1920
QY 638 AlaAlaHisLeuTyrLysHisSerGluAlaLeuGluValaLeuGlnAsnGlyIleLysSer 657
DB 1921 GCGGCCACCTGTGGAAGCAACATGAGGCGCTGAGAGCCCTGGAGACTTCATCAAGGCG 1980
QY 658 SerArgArgLeuGlnAsnPheCysArgAspPheGluLeuGlnLysValCysYrYrLeuPro 677
DB 1981 TCCGCGGCGCTGGAACATTTCTGCCGAGACTTGAGCTGAGAGAGGTGTTCCTGCGG 2040
QY 678 LeuAsnThrPheLeuLeuArgProLeuHisArgLeuMetHisYrLysGlnValIleGlnGlu 697
DB 2041 CTCAACACTTCTCTGCGCCCGCTGCAAGGCTCATGCACTTAAGCATGCTCTGAG 2100
QY 698 ArgLeuCysLysHisProProSerHisAlaAspPheArgAspCysArgAlaAlaLeu 717
DB 2101 AGCGTGGCAAGACCAACCCACAAACCAAGCCGACTTGAGGAGCTGCAAGCTGCGCTG 2160
QY 718 AlaGlnIleThrGlnMetValAlaGlnLeuHisGlyThrMetIleLysMetGluAsnPhe 737
DB 2161 GCGAGATTCACAGAGATGGTGGCAGAGCTGCAATGATCACTGATCAAGATGGAGAACTTC 2220
QY 738 GlnLysLeuHisGlnLeuLysLysAspLeuIleGlyIleAspAsnLeuValValProGly 757
DB 2221 CAGAGCTGCATAGACTCAAGAAAGATCTGATGGCATTAACAATCTTGATCCAGAGA 2280
QY 758 ArgGluPheIleArgLeuGlySerLeuSerLysLeuSerGlyLysGlyLeuGlnGlnArg 777
DB 2281 AGGAGATTATCCGCTGGGCAAGCTCCAGCAAGCTCCGAGAAAGGGGCTTCAGACGCG 2340
QY 778 MetPhePheLeuPheAsnAspValLeuLeuYrThrSerArgGlyLeuThrAlaSerAsn 797
DB 2341 ATGTTCTTCTGTTCAACGATGCTTGTGCTGATACAGCGCGGAGACTGAGAGCACTTAAT 2400
QY 798 GlnPheLysValHisGlyGlnLeuProLeuYrGlyMetThrIleGlnGlnSerGluAsp 817
DB 2401 CAGTTTAAATCCACGAGCAGCTCCACTTAATGCGATGACATCGAGAGAGTGAAGAG 2460
QY 818 GluTyrGlyValProHisCysLeuThrLeuArgGlyGlnArgGlnSerIleIleValAla 837
DB 2461 GAGTGGGGGTGTGCCACTGCTGACCTGCGGGGCAAGCGGCAATCTATCATGTGGCT 2520
QY 838 AlaSerSerArgSerGlnMetGlnLysTyrValGluAspIleGlnMetAlaIleAspLeu 857
DB 2521 GCCAGTTCCCGGTCAGAGATGGAATAATGATGAGAGACATCCAGATGGCATTTGACCTG 2580
QY 858 AlaGlnLysSerSerProAlaProGluPheLeuAlaSerSerProProAspAsnLys 877
DB 2581 GCAAGAAAGCAATGAGACCCACCCGAGCTCTTGCCAGAGAGCCCGCTTGACAAAG 2640
QY 878 SerProAspGluAlaThrAlaAlaAspGlnGlnSerGluAspAspLeuSerAlaSerArg 897

Db 2641 TCCCGGATGAGCCACAGCGGCTGACAGAGTACAGAGTACAGTCTCTCGG 2700

Qy 898 ThrSerLeuGluArgGlnAlaProH1bARG1yaenThrMetValH1bValCysTrpH1b 917

Db 2701 ACATGCTGAGAGCGCCAGGCGCCCTCAACCGTGGCAACAGATGTGTACAGTGTGGCAT 2760

Qy 918 ArgAsnThrSerValSerMetValaPhePheSerIleAlaValaGluAsnGlnLeuSerGly 937

Db 2761 CGCAGACACAGTGTCTTCATGTGTGAGACTTCACATTCCTGCTGGAGAACCATCTCTCGG 2820

Qy 938 AsnLeuLeuArgGlyPhePheLeuAsnSerAsnGlyIYrTrpGlnLysLeuTrpValaPheThr 957

Db 2821 AACTCTGCGAGAAATTCAGAAACAGCAATGCTGCGACAGACTGTGGGTATATTCACC 2880

Qy 958 AsnPheCysLeuPhePheTrpLysSerH1bGlnaPasnH1bProLeuAlaSerLeuPro 977

Db 2881 AACTCTGCTGCTGTCTTCTTACAAATACACACAGACAGATCCCTTGCCAGCTGCGCT 2940

Qy 978 LeuLeuGlyIYrTrpSerLeuThrIleProSerGluSerGluAsnIleGlnLysAspTrpVal 997

Db 2941 CTGCTGGGCTACTCCCTCAACATCCCTCTGAGTGTGAGAACATCCACAGAGACTATGTG 3000

Qy 998 PheLysLeuAlaPheLysSerH1bValIYrTrpPheArgAlaGluSerGlyTrpThrPhe 1017

Db 3001 TTCAAGCTGCATTTCAAGTCCCAACCTGATCTACTTCAGGCGTGAAGCAATACATTT 3060

Qy 1018 GluArgTrpMetGluValIleArgSerAlaThrSerSerAlaSerArgProH1bValLeu 1037

Db 3061 GAAAGGTGAGTGAAGTATCGAAGTGCACACACTCAGCTCTCGGGGCCACATCTTGG 3120

Qy 1038 SerH1bLysGluSerLeuValIYr 1045

Db 3121 AGTCACAAAGATCTCACCCTGAC 3144

RESULT 5
AK171713
LOCUS
DEFINITION
MUS musculus activated spleen cDNA, RIKEN full-length enriched library, clone:FE30006D05 product:FERM, Rhodsp and pleckstrin domain protein 2, full insert sequence.
AK171713
ACCESSION
AK171713.1 GI:74152707
VERSION
KEYWORDS
HNC; CAP trapper.
SOURCE
MUS musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Muridae; Murinae; Mus.

REFERENCE
AUTHORS
TITLE
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
PUBMED
10349636

REFERENCE
AUTHORS
TITLE
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
PUBMED
11042159

JOURNAL
PUBMED
AUTHORS
3
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Kono, H., Akiyama, J., Nishi, K., Kitsuana, T., Teshiro, H., Itoh, M., Suni, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kasai, K., Fujikawa, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Marahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system-384-format
Genome Res. 10 (11), 1757-1771 (2000)
PUBMED
11076861

REFERENCE
AUTHORS
TITLE
Kawai, J., Shingawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Aikawa, T., Hara, A., Fukunishi, Y., Kono, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamana, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischman, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pezole, G., Quackenbush, J., Schriml, L.M., Staudt, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Bares, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Butt, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombere, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seta, T., Shibata, Y., Storch, K.F., Suzuki, H., Toyo-oka, K., Wang, K.H., Weitz, C., Whitaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawai, J., Kohsaki, S. and Hayashizaki, Y.

CONSRMT
TITLE
RIKEN Genome Exploration Research Group Phase II Team and the PANTOM Consortium
Functional annotation of a full-length mouse cDNA collection
Nature 409 (6821), 685-690 (2001)
PUBMED
11217651

REFERENCE
AUTHORS
TITLE
CONSRMT
TITLE
Analyses of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
Nature 420 (6915), 563-573 (2002)
PUBMED
12466851

JOURNAL
PUBMED
AUTHORS
6
Carninci, P., Kasukawa, T., Katayama, S., Gough, J., Fritsch, M.C., Maeda, N., Oyama, R., Ravasi, T., Lenhard, B., Wells, C., Kodzius, R., Shimokawa, K., Bajic, V.B., Brenner, S.E., Batalov, S., Forrest, A.R., Zavolan, M., Davis, M.J., Wilming, L.G., Aidinis, V., Allen, J.E., Ambesi-Implombato, A., Apweiler, R., Attaliya, R.N., Bailey, T.L., Banai, M., Baxter, L., Beisel, K.W., Beresno, T., Bono, H., Chalk, A.M., Chu, K.P., Choudhary, V., Christofides, A., Clutterbuck, D.R., Crome, M.L., Dalla, E., Dalrymple, B.P., de Bono, B., Della Gatta, G., di Bernardo, D., Down, T., Engstrom, P., Fagiolini, M., Faulkner, G., Fletcher, C.F., Fukushima, T., Furuno, M., Furuki, S., Gariboldi, M., Georgi-Hemman, P., Gingeras, T.R., Gojobori, T., Green, R.E., Gustincich, S., Harbers, M., Hayashi, Y., Hensch, T.K., Hirokawa, N., Hill, D., Hummel, L., Iacono, M., Ikeo, K., Iwama, A., Ishikawa, T., Jakt, M., Kanapin, A., Katoh, M., Kawasawa, Y., Kelsio, J., Kitamura, H.,

Kitano, H., Kollias, G., Krishnan, S.P., Kruger, A., Kummerfeld, S.K., Kurochkin, I.V., Lareau, L.F., Lazarevic, D., Lipovich, L., Liu, J., Ljung, S., McWilliam, S., Madan Babu, M., Madera, M., Marchionni, L., Matsuda, H., Matsuzawa, S., Miki, H., Mignone, F., Miyake, S., Morris, K., Mottagui-Tabar, S., Mulder, N., Nakano, N., Nakauchi, H., Ng, P., Nilsson, R., Nishiguchi, S., Nishikawa, S., Nori, F., Ohara, O., Okazaki, Y., Orlando, V., Pang, K.C., Pavan, W.J., Pavoni, G., Pezold, G., Petrovsky, N., Piazza, S., Reed, J.F., Reid, J.F., Ring, B.Z., Ringwald, M., Roat, B., Ruan, Y., Salzberg, S.L., Sanderlin, A., Schneider, C., Schonbach, C., Sekiguchi, K., Sempke, C.A., Seno, S., Seppa, L., Sheng, Y., Shibata, Y., Shimada, H., Shimada, K., Silva, D., Sincalir, B., Sperling, S., Stupka, E., Sugita, K., Sultana, R., Takenaka, Y., Taki, K., Tammoja, K., Tan, S.L., Tang, S., Taylor, M.S., Teener, J., Teichmann, S.A., Ueda, H.R., van Nimwegen, E., Verardo, R., Wei, C.L., Yagi, K., Yamanihi, H., Zabarovsky, E., Zhu, S., Zimmer, A., Hide, W., Bult, C., Grimmond, S.M., Tesdale, R.D., Liu, E.T., Brusic, V., Quackenbush, J., Wahlestedt, C., Mattick, J.S., Hume, D.A., Kai, C., Sasaki, D., Tomaru, Y., Fukuda, S., Kanamori-Katayama, M., Suzuki, M., Aoki, J., Arakawa, T., Iida, J., Imamura, K., Itoh, M., Kato, T., Kawai, H., Kawagashira, N., Kawashima, T., Kojima, M., Kondo, S., Konno, H., Nakano, K., Ninomiya, N., Nishio, T., Okada, M., Pleassy, C., Shibata, K., Shiraki, T., Suzuki, H., Kawai, J., and Hayashizaki, Y. Watanuki, A., Okamura-Ono, Y., Suzuki, H., Kawai, J., and Hayashizaki, Y. FANTOM Consortium
The transcriptional landscape of the mammalian genome
Science 309 (5740), 1559-1563 (2005)
16141072

CONSRMT
JOURNAL
PUBMED
REFERENCE
AUTHORS

7
Kakumura, S., Tomaru, Y., Kasukawa, T., Waki, K., Nakanishi, M., Nakamura, M., Nishida, H., Yap, C.C., Suzuki, M., Kawai, J., Suzuki, H., Carninci, P., Hayashizaki, Y., Wells, C., Fritch, M., Ravasi, L., Pang, K.C., Hallinan, J., Mattick, J., Hume, D.A., Lipovich, L., Batalov, S., Engstrom, P.G., Mizuno, Y., Faghihi, M.A., Sanderlin, A., Chalk, A.M., Mottagui-Tabar, S., Liang, Z., Lenhard, B., and Wahlestedt, C.
Riken Genome Exploration Research Group
Antisense transcription in the mammalian transcriptome
Science 309 (5740), 1564-1566 (2005)
16141073

CONSRMT
JOURNAL
PUBMED
REFERENCE
AUTHORS

8 (bases 1 to 3743)
Arakawa, T., Carninci, P., Fukuda, S., Hachizume, W., Hayashida, K., Hori, F., Iida, J., Imamura, K., Imotani, K., Itoh, M., Kanagawa, S., Kawai, J., Kojima, M., Konno, H., Murata, M., Nakamura, M., Ninomiya, N., Nishiyori, H., Nomura, K., Ohno, M., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shiraki, T., Tagami, M., Tagami, Y., Waki, K., Watanuki, A., Muramatsu, M., and Hayashizaki, Y.
Direct Submission
Submitted (14-APR-2004) Yoshihide Hayashizaki. The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehico-cho, Tsurumi-ku, Yokohama, Kanagawa, 230-0045, Japan (E-mail: genome-res@gsr.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

TITLE
JOURNAL

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Tissues were provided by Dr. John Todd (Dept. of Medical Genetics Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome Trust/MRC building Addenbrookes Hospital Cambridge) whose assistance we gratefully acknowledge.
Please visit our web site for further details.
URL: http://genome.gsc.riken.jp/
URL: http://fantom.gsc.riken.jp/
Location/Qualifiers

COMMENT

FEATURES
SOURCE

1. 3743
/organism="Mus musculus"
/mol_type="mRNA"
/strain="M01"
/db_xref="PANTOM:DB:P830006D05"
/db_xref="taxon:10090"

CDS

/clone="P830006D05"
/csize type="activated spleen"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
136..3333
/note="unnamed protein product; PERM, RhoGAP and pleckstrin domain protein 2 (MGD MGI:2385126 GB|BC009153, evidence: BLASTN, 99%, match=3696)
putative"
/protein_id="BAE42626.1"
/db_xref="GI:74152708"
R1RVK1LDSVLFEDIPKDDGVLLTQWKNLLECDYFGLFKNVOSYWMLEPM KPIIRKPKVAVLAVKFPDPGLOEETRYLFLQRLDLLESLTCTANTA ALLSHLQSEIDYDETLDEHLEKANEYLPNQKSLKILDEHSHRTQTPAESDFQ VLEIARKEIYGRFMAADREGTKNLINLVSHGVIVPGTTTINFMNPKSLSPK RKPFLKHPBVGHPVDTLPELPSRDECKNWKICVETHTTFRSLSDQPKPAKVF FSRGSSFRYSRTQKQVVDVYDGMKRIPEYRHSKRTSLHALTVDPKQSVSFTD

Alignment Scores:

Pred. No.:	5,86e-241	Length:	3743
Score:	2916.00	Matches:	562
Percent Similarity:	70.1%	Conservative:	179
Best Local Similarity:	53.2%	Mismatch:	268
Query Match:	53.4%	Indels:	48
DB:	6	Gaps:	9

US-09-555-342B-2 (1-1045) x AKT17173 (1-3743)

QY	1	MetGlyGluIleGlu-----GlnAyrProThrProGlySerAtrGluGlyAla	16
DB	136	ATGGAGAGATAGAGAAACATGACAGAGCTTTTCCAACTTCAGAGGACACCTGGGGTGC	195
QY	17	ProGluAnsSerGlyIleSerThrLeuGluAyrGlyGlnAyrProProProThrProSer	36
DB	196	CAGACCGCATTTGAGAGGATACCTGAGACCCAGAGACGCTGTCCACAGATGAC	235
QY	37	GlyLeuLeuValSerIleLeuValIleGlnMetLeuAspAyrThrGlnGluAlaPheGluVal	56
DB	256	GAGAGACATGCGGATCAGAGAGGAGGAGGCTGTGACGACCTGTAGGATTATGACATC	315
QY	57	ProGlnAyrAlaProGlyValValLeuLeuAspAlaValCysAsnHsIleLeuAnLeuVal	76
DB	316	GAGCCAAAGTGTACGGCGGCTTACCTGACTCAAGTGTGGAAGCATTTTGAATTCATTT	375
QY	77	GluGlyAspTyrPheGlyLeuGluPheProAspHsIleValSerIleThrValTTPLeuAsp	96
DB	376	GAATGTGACTACTTTTGTGTGAGTTCCAGAAATGTCAGTCCATCTGATTTGGCTTGA	435
QY	97	LeuLeuLeuProIleValValGlnIleAyrAyrProGlyHsIleValValValPheVal	116
DB	436	CGATGAAACCATCATCTTAGGCAAGTACGAAACCAAAATGCGGTTCCTGCGCTGCA	495
QY	117	ValLeuPhePheProProAspHsIleThrGlnLeuGlnGluIleuThrAyrGlyrLeuPhe	136
DB	496	GTAATAATTTTCCCGCTGACCTGTGTGAGTGCAGAAAGATGACCAAGATGACCTGTTT	555
QY	137	AlaLeuGlnValValGlnAspLeuAlaGlnGlyAyrGlnLeuThrCysAsnAyrThrSerAla	156
DB	556	GCTTGCATCTCAAGAGAGACTCTCTGAGAAACGCTGACTGACGCGCAACATGCA	615
QY	157	AlaLeuLeuIleSerHsIleValGlnSerGluIleGlyAyrPheAspGluAlaLeuAsp	176
DB	616	GCCCTTCTCATATCCACCTTGTGAGGTGGAATGCGAGATATATGATGAACCTTGAT	675
QY	177	AyrGlnHsIleLeuAlaLeuAnLysTyrIleProGlnGlnAspAlaLeuGluAspGlyIle	196
DB	676	CGAAGACACCTCAAGACCATGATGATCTTCCCAACAGGAAATCTTCAAGAAAGATA	735
QY	197	ValGluPheHsIleAsnHsIleIleGlyGlnThrProAlaGluSerAyrPheGlnLeuLeu	216
DB	736	CTAGACTTCATCAGAGGACACGCGGACAGACTCTGCAAGTACAGATTTCCAGGTGCTT	795

QY 217 GluIleAlaArgXArgLeuGluIleuMetTyrGlyIleArgLeuHisProAlaIleuAspArgGlu 236
Db 796 GAGATTGCAAGAACCTGGAAATGTAATGGCATCAGGTTCCGATGGCTTTCAGACAGAA 855
QY 237 GlyThrIleuValAsnLeuAlaValAlaAsnThrGlyIleLeuValPheGlnGlyPheThr 256
Db 856 GGGACCAAGATTAACTAGCAGTTCTCCATGGGGTGTCTGTGTTCCAGGGTACACCC 915
QY 257 LeuIleAsnAlaPheAsnThrAlaIleuValArgIleuSerPheIleuArgIleuArgPhe 276
Db 916 AAAATCAATACCTTAACCTGCTCAAGCTCCAAAGCTCCAAAGCTTAAGCTTAAGAAAGGTTT 975
QY 277 LeuIleLeuLeuArgProAspAlaAsnSerAlaIleuArgIleuAspThrLeuGlnPheIleuMet 296
Db 976 CTTATCAAACTCCACCTGAGGTCCATGGGCTTCACAGACACCTTAAGTCTTGTGG 1035
QY 297 AlaSerArgAspPheCysIleuSerPheThrIleuValCysValGlnHisIleAsnIlePhePhe 316
Db 1036 GGTACACAGATGATGATGAAGACTTCTGGAAGATATGTGTGAGATACCATCTTTT 1095
QY 317 ArgLeuPheGlnGluProIleuArgProIleuArgProIleuPheSerArgIleuSer 336
Db 1096 AGACTCTGTGACCACTTAAGCCAAAGGCAAGGCTGCTCTTCCAGCCAGGCTCTCC 1155
QY 337 PheArgPheSerGlyArgThrGlnIleuValIleuAspIleuValIleuGlnGlyIleuHis 356
Db 1156 TTCAGATTCAGTGGAAAGTCAAGAACTCAAGAACTAGATTAATGTAAGACGGTGGAAATG 1215
QY 357 LeuIleValGlnPheGluArgIleuHisIleuSerIleuHisIleuSerIleuValIleuSer 376
Db 1216 AAGGAATTCATAGAAAGACGACAGTAAAGCTCGG----- 1254
QY 377 GlnProThrGluLeuAsnSerGluValIleuGlnIleuSerGlnIleuSerThrSerIleuThr 396
Db 1355 -----ACATCTCTTCACTCTGACTGTAGATCTGCTTAACAGAGGCTCTTCAAC 1308
QY 397 PheGlyGlnGluIleuArgIleuSerProIleuGlnIleuSerCysArgArgGlyIleuGluProIleu 416
Db 1309 GATGGCTTGAGAACTTCTGCC-----TCC 1332
QY 417 ValSerAlaGlyGluProGlyIleuSerHisIleuProSerProAlaProArgArgSerProAlaGly 436
Db 1333 CTGTCTTCAGCAAAATGTCTCTTTTATTCACCCCTAGTCTTCTTGTGCTCTCTGCT 1392
QY 437 AsnIleuSerGlnAlaAspGlyAlaIleuSerAlaProThrGlnGluGlnGluValIleuValys 456
Db 1393 CTGCCCAATTGTAAGACAGACGACGCTCTTGTGTGATCCGACAGCTCCCTCATTCAG 1452
QY 457 Asp-----ArgThrGlnGln 461
Db 1453 AGCACACAGACAGAGACAGACGACCATCATCTGACGGCCCAACACACAGATCG 1512
QY 462 SerIleuS---ProGlnProPro-----GlnProSerThrGlySerIleuThrGlySerPro 478
Db 1513 GCCCATCTCCCTGAGCCCCCGGTGTCTCGGCTGTCCAGGCTTTTCATGATATGTCCT 1572
QY 479 HisIleuSerGluLeuSerValAsnSerGlnGlyIleuValAlaProAlaAsnValIleu 498
Db 1573 CAGCTTCTCCCTCCAGCTGTAAGCACTGACCTGTGCTCT-----GAGCTT 1623
QY 499 SerProAsnLeuSerProAspThrIleuValIleuSerProLeuIleuSerProLeuLeu 518
Db 1624 CAGCCCGGCTGAGTACAGCTGACGAGGTCATCCCTGTGCTACAGCCCTGTGTCTGCT 1683
QY 519 AspGlnAlaCysProArgThrAspArgGluIleuArgGlyIleuArgIleuArgPheProThr 538
Db 1684 GCGCTGTGACAGCAGATGGAACAACAAAGACAGAAACACAAACATGTCAGAA 1743
QY 539 AspIleuValIleuThrIleuAlaIleuGluValIleuThrIleuGlnIleuArgIleuValAsp 558
Db 1744 GATGAGGCTTAATTTATAGCCAAAGAGATTTCTGCTACAGAAACCTATCTGAAGAT 1803
QY 559 LeuGluValIleuThrSerIleuThrPheGlnSerThrValSerIleuValAspAlaIleuMetProGlu 578

Db 1804 TTAAGATTATCAGCTGTGCTTCCGACGCTGGTANTCAAGAGAGGACCATGCTGCA 1863
QY 579 AlaIleuIleuSerIleuIleuPheProAsnPheGluProIleuHisIleuPheHisThrAsnPhe 598
Db 1864 GCCCTGATGAGCCCTGCTTTTCTCAACATTAATCAGCTACAGATTCAGAGTTCACAGAGGCTTC 1923
QY 599 LeuIleuGlnIleuGlnIleuValIleuThrIleuGlnIleuValIleuArgIleuValIleuArg 618
Db 1924 CTCACGAGGTGAACAGAGCTGACCTGAGGAGGCGCTCCAGTGCCTTAATA 1983
QY 619 ---AspTyrGlnIleuArgIleuValIleuAspValIleuSerIleuValIleuGlnIleuValIleu 637
Db 1984 GGTATTCACCAAGCATGAGGATGATCTCTCCAGAAACATGCTTAAGAAATT 2043
QY 638 AlaAlaHisIleuThrIleuHisIleuSerGluAlaIleuGlnIleuValIleuGlnIleuValIleu 657
Db 2044 ACTAGCTACTTCCAAAGACAGATGAGGCTCTTAACAGAACTGAAAGGCCACAAACAC 2103
QY 658 SerArgArgLeuGluIleuAsnPheCysArgAspPheGluIleuGlnIleuValIleuPro 677
Db 2104 TGTAAAGCTGAGAGGACGTCTTAACAAAGTTGAGCTCAAAAGTGTGCTACCTGCT 2163
QY 678 LeuAsnThrPheLeuLeuArgProIleuHisIleuGluMetHisIleuValIleuGlnIleuGlu 697
Db 2164 CTCAAACATTCCTGTGTGAACCCGTCACAGGCTAGTCCATCCCTGTGCTGAGC 2223
QY 698 ArgLeuCysIleuHisIleuProSerHisIleuAspPheArgAspCysArgAlaIleu 717
Db 2224 CGCGTGTGTCTACTACTCTCTGAGCAGCGCACTATGTCCGCTGACAGAGGCACTG 2283
QY 718 AlaGluIleuThrGluIleuValIleuGlnIleuHisIleuMetIleuMetGluAsnPhe 737
Db 2284 AAGGCATCAAGAAAGTATACCAAGGCTCAAGAAAGCTTACCCGCTGAAACCTA 2343
QY 738 GlnIleuSerHisIleuLeuIleuValIleuAspIleuIleuGlyIleuAsnIleuValIleuProGly 757
Db 2344 CAGAAATTGACGAGAGCTACAGAGAGCTGCTGCTGATGAAGAACTCATGCTCTGCTG 2403
QY 758 ArgGluPheIleuArgLeuGlyIleuSerIleuSerIleuSerIleuGlyIleuGlnIleuArg 777
Db 2404 AGGAGATTATCCGTGAGGCTGCTGCACAAGCTCAACAAAGAGGCTGCAACAGAG 2463
QY 778 MetPhePheIleuPheAsnAspValIleuIleuThrIleuSerArgIleuThrAlaSerAsn 797
Db 2464 ATGTTTCTGTTCTCAGATATGTACTGTATCAAGCAAAAGTGTACAGAGCCAGT 2523
QY 798 GlnPheIleuValHisIleuGlnIleuProIleuTyrGlyIleuMetThrIleuGlnIleuSerGluAsp 817
Db 2524 CATTTCCGAGTCCGTCGCTCTTCCACTCCGTGGCATGCTGTGAAGAAAGTGAAGAT 2583
QY 818 GluThrGlyValProHisCysLeuThrIleuArgIleuArgIleuArgIleuValIleuValAla 837
Db 2584 GAATGTCTGTTCTTCAATTGCTTCACTATCATGACAGTCAAGAAACATATGTGTGAGCA 2643
QY 838 AlaSerSerArgSerGluMetGluIleuThrValIleuIleuIleuIleuMetAlaIleuAspLeu 857
Db 2644 GCCAGCACTCGCTAAGAAAGAAAGTGAAGAGAGCTGAATGCACATCAACAA-- 2700
QY 858 AlaGluIleuSerSerSerProAlaProGluIleuLeuAlaSerSerProProAspAsnIleu 877
Db 2701 GCAGCCAAAGATATTCGGATGTACACCCCAATGCTGTGGAGGCGCGGTATATCTGCT 2760
QY 878 SerProAspGluAlaThrAlaIleuPheGlnIleuSerGluAspAspLeuSerIleuSerArg 897
Db 2761 ACCCTGATGATCTGTGAAGTCTCTGTGAAGAAATCAGAAAGTGTGAGAGAAACCG 2820
QY 898 ThrSerIleuGluArgGlnAlaProHisArgGlyIleuAsnThrMetValHisIleuValCysThrHis 917
Db 2821 GGTCTCCGTGAGGAGAACAGCAGACCGGGCCAAATACAAATGCAATGATGTGTGCTGATC 2880
QY 918 ArgAsnThrSerValSerMetValAspPheSerIleuAlaValGluAsnGlnIleuSerGly 937

Db 2881 CGTAAATCAAGTGTGTCCAGACGACACAGTCAGCTGTGGAACACGCTTTCAGGA 2940

Qy 938 AsnLeuLeuArgLysPheLysAsnSerAsnGlyTTPglnLysLeuTTPValPheThr 957

Db 2941 TATTCGTCGAGAAAGTTCAAGAACGTAATGCTGGAGAAAGCTCTGGGTGCTTTTACC 3000

Qy 958 AsnPheCysLeuPhePheTyrLysSerHisIleAspAsnHisProLeuValAsnLeuPro 977

Db 3001 AACTTCGCTGTTCTTCTTAATAAACATCAGATGATGATACCCCTGGCCAGGCTCCCA 3060

Qy 978 LeuLeuLysTyrSerLeuThrIleProSerGlnSerGlnLysLeuIleGlnLysAspTyrVal 997

Db 3061 CTATTTGGGTATACACCGAGGCTCCACAGGAGGCTACACGACATCCAAAGACTATGTC 3120

Qy 998 PheLysLeuHisPheLysSerHisValTyrTyrPheArgIleGlnSerGlnTyrThrPhe 1017

Db 3121 TTCAAGCTCCAAATTCAAATCTCACGCTTACTCTTCCGGGCTAGACAAAGTACACATTT 3180

Qy 1018 GlnArgTTPMetGlnValIleArgSerAlaThrSerSerAlaSerArgPro 1034

Db 3181 GAAGAAGTGATGACGTCATCAAAAGGCGCAGAGCTCACAGGAGAGACCC 3231

RESULT 6

AKO50860 3874 bp mRNA linear HTC 02-SEP-2005

LOCUS AKO50860.1 GI:26094164

DEFINITION Mus musculus 9 days embryo whole body cDNA, RIKEN full-length enriched library, clone:DO30026M03 product:weakly similar to CDEP (Homo sapiens), full insert sequence.

ACCESSION AKO50860

VERSION AKO50860.1 GI:26094164

KEYWORDS HTC; CAP trapper.

SOURCE Mus musculus (house mouse)

ORGANISM Eukaryotes; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Muridae; Murinae; Mus.

REFERENCE 1 Carninci, P. and Hayashizaki, Y. High-efficiency full-length cDNA cloning Meth. Enzymol. 303, 19-44 (1999) 10349636

REFERENCE 2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, H., Konno, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000) 11042155

JOURNAL PUBMED 11042155

REFERENCE 3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kiteunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishiguchi, T., Harada, K., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kaishoji, K., Fujiwaka, S., Inoue, K., Togawa, Y., Itawa, M., Onari, E., Matsubara, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsunaga, S., Kawai, D., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multiplexillary sequencer Genome Res. 10 (11), 1757-1771 (2000) 11076861

JOURNAL PUBMED 11076861

REFERENCE 4 The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium. Functional annotation of a full-length mouse cDNA collection Nature 409, 685-690 (2001)

JOURNAL REFERENCE 5 The FANTOM Consortium, the RIKEN Genome Exploration Research Group Phase I and II Team. Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs Nature 420, 563-573 (2002)

JOURNAL REFERENCE 6 RIKEN Genome Exploration Research Group, Genome Science Group

TITLE (Genome Network Core Team) and the FANTOM Consortium. Antisense Transcription in the Mammalian Transcriptome Science 309, 1564-1566 (2005)

JOURNAL REFERENCE 7 The FANTOM Consortium, Riken Genome Exploration Research Group and Genome Science Group (Genome Network Project Core Group). The transcriptional landscape of the mammalian genome Science 309, 1559-1563 (2005)

JOURNAL REFERENCE 8 (bases 1 to 3874)

AUTHORS Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, P., Imotani, K., Ishii, Y., Itoh, M., Kagawa, T., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M., and Hayashizaki, Y.

TITLE Direct Submission

JOURNAL Submitted (16-JUL-2001) Yoshinori Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration and Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa, 230-0045, Japan (E-mail: genome-resgsc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details. URL: http://genome.gsc.riken.jp/ URL: http://fantom.gsc.riken.jp/ Location/Qualifiers

FEATURES

source

1..3874

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="FANTOM_DB:DO30026M03"

/db_xref="taxon:10090"

/clone="DO30026M03"

/issue_type="whole body"

/clone_lib="RIKEN full-length enriched mouse cDNA library"

/dev_stage="9 days embryo"

1..3874

/note="putative weakly similar to CDEP (Homo sapiens) (SPTR|Q9Y4F1, evidence: FASTV, 53.7%ID, 99.7%length, match=1125)"

ORIGIN

Alignment Scores:

Pred. No.: 3,03e-239 Length: 3874

Score: 2896.50 Matches: 565

Percent Similarity: 70.3% Conservative: 181

Best Local Similarity: 53.3% Mismatches: 261

Query Match: 53.0% Indels: 55

DB: 6 Gaps: 12

US-09-555-342B-2 (1-1045) x AKO50860 (1-3874)

Qy 1 MetGlnIleGlu-----GlnArgProThrProGlySerArgLeuGlyAla 16

Db 70 ATGGAGAGAGTAAAGGACATACAGAGCTTTACCACTTCAGGAGACCGCTTGGGTGCC 129

Qy 17 ProGlnAsnSerGlyLeuSerThrLeuGlnArgGlyGlnLysProProProSer 36

Db 130 CACACCGCATTTGGATGAGTACCTCGGAGCCAGACAGAGCGCTGCACCCAGAGTCAG 189

Qy 37 GlyLysLeuValSerIleLysIleGlnMetLeuAspThrThrGlnGluAlaPheGluVal 56

Db	190	GAGAAAGCACAATCGGCATTCAGAGAGAAAGCTGCTGAGACGCACTGTAAAGTTATTGTGACATC	249
Oy	57	ProGlnAlaArgAlaProGlyLysValLeuLeuAspAlaValCysAsnHisLeuAsnLeuVal	76
Db	250	GAGCGGAAAGTGTGACGGGACAGGCTTTACAGCTCAAGTGTGGAAAGATTGAACTTCATT	309
Oy	77	GluGlyAspArgTyrPheGlyLeuGluPheProAspHisLysLysValLeuValTrpLeuAsp	96
Db	310	GAATGTGACTACTTGTGGTGTGGAGTTCAAGAAATGTGCATGCTCTCACTGAATTTGGTCTGAA	369
Oy	97	LeuLeuLysProGluValValysGlnIleAspArgProLysHisValValLysPheVal	116
Db	370	CCTAAGAAACCATCATATTGAGCAAGACAAAGAACCAAAATGCGGTGCTCGCCCTGGCA	429
Oy	117	ValLysPhePheProProAspHisTrpGlnIleGluGlnGluGluCysLeuHisArgTyrLysPhe	136
Db	430	GTAATAATTTTCCCGCCTGACCCCTGGTCAGCTGCACAAAGAAAGATGACAAAGGACCTGTT	489
Oy	137	AlaLeuGlnValLysGlnAspLeuAlaGlnGlyArgLeuThrCysAsnAspTrpSerAla	156
Db	490	GCCCTTGCAACTCAAGAGACCTCTCGAGGAAAGCGCTGACCTGACCGGCAACCTGGAT	549
Oy	157	AlaLeuLeuIleSerHisLysValGlnSerGluIleGlyAspPheAspGluAlaLeuAsp	176
Db	550	GCCCTTCTCATATCCCACTTGTGCACTGGGAAATGGAGATTATATGATAAACCTCGAAT	609
Oy	177	ArgGlnHisLeuAlaLysAsnLysTyrLysProGlnGlnAspAlaLeuGluAspLysIle	196
Db	610	CGAGAACACTCAAAAGCCAAATGATTTCTTGCCCAACCAAGAGAAATCTTAGAAAGATA	669
Oy	197	ValGluPheHisValAsnHisArgLeuGlnTrpAlaGluSerAspPheGlnLeuLeu	216
Db	670	CTAGACTTCCATCCAGAGGCAACGGGCGAGACTCCGCGAGAGTCAGATTCCAGGTCGCTT	729
Oy	217	GluIleAlaArgArgLeuGlnMetTyrGlyIleArgLeuHisProAlaLysAspArgGlu	236
Db	730	GAGATTGGCAAGGAGCTGGAAATGTATGTGCATCAGAGTTTCACATGGCTTCTGACAGAA	789
Oy	237	GlyThrLysIleAsnLeuAlaValAlaAsnTrpGlyIleLeuValPheGlnGlyPheThr	256
Db	790	GGGACCAAGATTAACTTCAAGATTTCTCAAGAGGAGTCTGTGGTTTCCAGGGTACCAC	849
Oy	257	LysIleAsnAlaPheAsnTrpAlaLysValArgLysLeuSerPheLysArgLysArgPhe	276
Db	850	AAATCAATACCTTCACTGGTCCAAAGTCCGTAACTAAAGCTTCAAGAGAAAGGTTT	909
Oy	277	LeuIleLysLeuAspProAspAlaAsnSerAlaTyrGlnAspTrpLeuGlnPheLeuMet	296
Db	910	CTTATCAAACTCCACCTGAGGTGTCATGGGCGCTTCAAGAGACCTTAAAGTTCTTGTTG	969
Oy	297	AlaSerAspAspPheCysLysSerPheTrpLysIleCysValGlnHisIleAlaPhePhe	316
Db	970	GGTACCAAGAGAAATGTAAAGAACTTGTGAAAGATATGTGTGAGATCAACAATCTTTT	102
Oy	317	ArgLysPheGluGlnProLysArgProLysProValLeuPheSerArgLysSer	336
Db	1030	AGACTCTGACACAGCTTAAGCCAAAGGAGGAGCGCTGTTCTTCAAGCGGAGCTCTCC	108
Oy	337	PheArgPheSerGlyArgThrGlnLysGlnValLeuAspTyrValLysGluGlyHis	356
Db	1090	TTCAAGATACAGTGGAAAGAACTCAAGAAACAATGATGATTATGTCAAAGACGCTGGAAG	114
Oy	357	LysLysValGlnPheGluArgLysHisAspLysValLeuSerIleArgSerLeuAlaSer	376
Db	1150	AAAGAAATTCATCAACAAAGACGCAACAGTAAGACTCGG-----	1188
Oy	377	GlnProThrGlnLeuAsnSerGluValLeuGlnGlnSerGlnGlnSerHisSerLeuThr	396
Db	1189	-----ACATCTCTTCACTCTCGACTGTGAATGTGCGCTAAACAGAGCGTCTTCAAC	124
Oy	397	PheGlyGluGlyAlaGluSerProGlyGlyGlnSerCysArgArgGlyLysGluProLys	416

Db	1243	GATGGCTTGAGAACTTCTGCC-----TCC	1268
Qy	417	ValSerAlaGlyGluProGlySerHisProSerProAlaProArgArgSerProAlaGly	436
Db	1267	CTGTCCTTCAGGAAGATGCTCTCTTTATTCACCCCTTAAGTCTTCTGTTCCCTCTGGT	1326
Qy	437	AsnIysGlnAlaAspGlyAlaIAserAlaProThrGluGluGluGluValValIys	456
Db	1337	CTGCCCAATTGAAAGACAGACAGACGTCCTTGAGATCCGACAGGCTCCCGTCATCAAG	1386
Qy	457	Asp-----ArgThrGlnGln	461
Db	1387	AGCACAGACAGACAGAGAGACAGCGACATCATCATCTTACCGGCCACACACATCGC	1446
Qy	462	SerIys---ProGlnProPro-----GlnProSerThrGlySerLeuThrGlySerPro	478
Db	1447	GCCCATCTCCCTGGACCCCCCGGTGTCGGCGTGGTCAGGCTTTTCTATGATGATGCTCT	1506
Qy	479	HisLeuSerGluLeuSerValAsnSerGlnGlyValAlaIAlaProAlaAsnValThrIleu	498
Db	1507	CAGCCTTCTCCCTCCAGCCTTAAGAGCCACCTGACGCTGTCGCTT-----GAGCTT	1557
Qy	499	SerProAsnLeuSerProAspThrIleGlnIleAserProIleIleSerProIleuAsn	518
Db	1558	CAGCGCCGCGTGAAGTACAGTCAAGCAAGGTCATCTCCCTGTGCTCAAGCCTCTGTCTAGT	1617
Qy	519	AspGlnAlaCysProArgThrAspAspGluAspGluIleArgArgIysArgPheProThr	538
Db	1618	GGCGCTGGTACAGCGAGATGGATGGACACCAAGAGAGAGAAACACAGCATCGCAGAA	1677
Qy	539	AspIysValaIysPheIleAlaIysGluValSerThrThrGluArgThrTyrLeuIysAsp	558
Db	1678	GATAGAGGCTTTTCATGATGACCAAGAGATGTTCTCCCTACAGAACCAACTATCTAAGAGAT	1737
Qy	559	LeuGluValIleThrSerThrPheGlnSerThrValSerIysGluAspAlaMetProGlu	578
Db	1738	TTAGAGATTACACTGTGTGGTTCGCGAGCTGGATCAAGAGAGAGGCGATCGCTCA	1797
Qy	579	AlaLeuIysSerLeuIlePheProAsnPheGluProLeuHisIysPheHisThrAsnPhe	598
Db	1798	GGCTGTAGAGGCGCTTCCTTTCTCCAACTTATGATCAAGTCAAGATTCACAGAGGCTTC	1857
Qy	599	LeuIysGluIleGluGlnArgLeuAlaLeuTyrGluIleArgSerAsnAlaGlnIleArg	618
Db	1858	CTTCACAGAGGTGGAACAGAGGCTGGCACTCTGGGAGAGGCGCTCCAGAGCCCATTAATA	1917
Qy	619	---AspTyrGlnArgIleGlyAspValMetLeuIysValAsnIleGlnIleMetLeuHisIleu	637
Db	1918	GGTATGATCCACGCAATTCGGGAGATCTCTCTCCAGGAACAGCGCTCAAGTTAAAGAAATT	1977
Qy	638	AlaAlaHisLeuTyrIleHisSerGluAlaLeuGluAlaLeuGluValAsnGlyIleIysSer	657
Db	1978	ACTAGCTACTTCCAAAGACACAGATGAGTCTTAAACAGACTCGAAAAAGCCACAAAACAC	2037
Qy	658	SerArgArgLeuGluAsnPheCysArgAspPheGluLeuGlnIysValCysTyrLeuPro	677
Db	2038	TGTAAAAAGCTGAGGACAGCTCAAAAGAGTTTAGCTCCAAAAGTCTGCTACTGCTCT	2097
Qy	678	LeuAsnThrPheLeuLeuArgProLeuHisArgLeuMetHisTyrTyrGlnValLeuGlu	697
Db	2098	CTCAACACATTTCTCGCTGAAGCCCGCTCCAGAGGCTAGTGCATACCTCGCTGCTAGAC	2157
Qy	698	ArgLeuCysIysHisIleProProSerHisAlaAspPheArgAspCysArgAlaIleu	717
Db	2158	CGGCTGTGTGCTACTACTCTCTCTGGGCAACCGCACTATGCGCACTGCAAGAGCACTG	2217
Qy	718	AlaGluIleThrGluMetValAlaGlnLeuHisGlyThrMetIleIysMetGluAsnPhe	737
Db	2218	AAGGCGATTCACAGAAAGTGAACCAACGAGCTCCAGCAAAAGCTTATCCGCGTGAATACTTA	2277
Qy	738	GlnIysIleHisIleGluLeuIysIleAspLeuIleGlyIleAspAsnIleValIleProGly	757
Db	2278	CAGAAATTTGACGAGACTACACGCAAGACTGGTGTGCTGTGAAGAACTTCATATGCTCTGGG	2337

QY 758 ArgGluPheIleAArgLeuGlySerLeuSerLeuSerGlyLeuGlyLeuGlnArg 777
Db 2338 AGGAGATTATTCCTGAGGGCTGCTGCACAGACTCACCAAGAGGGCTGCACACAGG 2397
QY 778 MetPhePheLeuPheAenAapValLeuLeuTyThrSerArgGlyLeuThrAlaSerAsn 797
Db 2298 ATGTTTCTTCTGCTTCAGATATGCTGCTGTAACAGCAAAAGATGTCACAGAGCCAGT 2457
QY 798 GlnPheLeuValHISGlyGlnLeuProLeuTyTrGlyMetThrIleGlnGlnSerGlnAsp 817
Db 2458 CATTTCGGAGTCCGTGCTTCCTTCACCTCGTGCGCATGCTGTGAGAGAAAGAGAAAT 2517
QY 818 GlnTyPrlGlyValProHISCyLeuLeuThrLeuArgGlyGlnArgGlnSerIleIleValAla 837
Db 2518 GAATGGCTGTCTCTCTCATGCTTTCACCATCATACAGCTCAGAAACAAATTTGCTAGCA 2577
QY 838 AlaSerSerArgSerGlnMetGlnLeuTyTrValGlnLysPrlIleGlnMetAlaIleAspLeu 857
Db 2578 GCCAGCACTCGGCTAGAAAGAAAGATGATGACAGACCTGATGACAGATCCAA--- 2634
QY 858 AlaGlnLysSerSerSerProAlaProGlnPheLeuAlaSerSerProProAspLys 877
Db 2635 GCAGCCAAAGACTATCGGTGACTCACCCAGTGTCTGTGGAGGCCGCGGTATATCTGT 2694
QY 878 SerPro-----AapGlnAlaThrAlaAlaSerGlnGlnSerGlnAspLeuSer 894
Db 2695 ACCCTTGATCTTGTGATGAGTCTCTG-----GAAGATGAGAAAGATGCTCG-AGG 2747
QY 895 AlaSerArgThrSerLeuGlnArgGlnAlaProHISArgGly--AanThrMetValHIS 913
Db 2748 AAAGCCGGGGCTGCTGAGAGGGAGAACAGCCAGACCGGGCCCAATATCAACAATGCAT 2807
QY 914 ValCyThrPrlHISArgAenThrSerValSerMetValAspPheSerIleAlaValGlnAsn 933
Db 2808 GTGGCTGTGTAACCTGATATACAGGTGTGCCAGACACCAACATGTGACGTGTGAGAAC 2867
QY 934 GlnLeuSerGlnAenLeuLeuArgLysPheLysAsnSerAsnGlyTrpGlnLysLeuTyTr 953
Db 2868 CAGCTTCAGAGATATCTGCTGAGAAAGTTCAGAAACAGTAAAGTCTGCGACAGAGCTTGG 2927
QY 954 ValValPheThrAenPheCyLeuPhePheTyLysSerHISGlnAspAsnHISProLeu 973
Db 2928 GTGGCTTTTACCACTTCTGCTGCTTCTCTATAAACAATCAGATGACTACCCCTG 2987
QY 974 AlaSerLeuProLeuLeuGlyTyTrSerLeuThrIleProSerGlnSerGlnAsnIleGln 993
Db 2988 GCCAGCTCCCACTATTTGGGTTCAGGCTAGCTCCCAAGGAGGCTGACAGATCCAC 3047
QY 994 LysAspTyTrValPheLysLeuHISPheLysSerHISValTyTrPheArgAlaGlnSer 1013
Db 3048 AAAGACTATGCTCTTCAGAGCTCCATTCATATCTACCTTCTTCCGGGCTGAGAGC 3107
QY 1014 GlnTyThrPheGlnArgTrpMetGlnValIleArgSerAlaThrSerSerAlaSerArg 1033
Db 3108 AAGTACACATTTGAAAGTGATGATGACATCAAAAGGCCAGCACTCACCAGGAGAGA 3167
QY 1034 Pro 1034
Db 3168 CCC 3170

RESULT 7
AY415515 3165 bp DNA linear GSS 17-DEC-2003
LOCUS DEFINITION Homo sapiens FARP2 gene, VIRIDAL TRANSCRIPT, partial sequence,
AY415515 genomic survey sequence.
VERSION AY415515.1 GI:39771474
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

REFERENCE 1 (bases 1 to 3165)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civeille,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 3165)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civeille,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source location/Qualifiers
1..3165
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1..>3165
/gene="FARP2"
/locus_tag="HCMS583"
ORIGIN
Alignment Scores:
Pred. No.: 2,92e-189 Length: 3165
Score: 2318.50 Matches: 486
Percent Similarity: 59.9% Conservative: 149
Best Local Similarity: 45.8% Mismatches: 368
Query Match: 42.4% Indels: 57
Gaps: 14
US-09-555-342B-2 (1-1045) x AY415515 (1-3165)
QY 1 MetGlyGlnIleGlnGlnArg-----ProThrProGlySerArgLeuGlnVal 16
Db 1 ATGGGGAGATGAAGAAACATACAGAGTCTGTCAGACTGACAGGATGCGCTTGGGTGCC 60
QY 17 ProGlnAsnSerGlyIleSerThrLeuGlnArgGlyGlnLysAspProProThrProSer 36
Db 61 CAGACCCCTGTGGAGATTCAGACCTTGAACCTGGGCAAGCTCTTGGCCAGAAATGCAA 120
QY 37 GlnLysLeuValSerIleLysIleGlnMetLeuAspAspThrGlnLysAlaPheGlnVal 56
Db 121 GAGAAAGACCTGCACCTCAGAGTAAAGCTGCTGACAAACCATGGAATATTTGACATT 180
QY 57 ProGlnArgAlaProGlyLysValIleLeuAspAlaValCysAsnHISLeuAenLeuVal 76
Db 181 GAGCTTAATTCGATGCGCAGGATTAATCTGACACAAAGTGTGGAAGCGTTTAAACCTGTA 240
QY 77 GlnGlyAspTyTrPheGlyLeuGlnLysPheProAspHISLysValSerValTrpLeuAsp 96
Db 241 GAATGTGACTACTCTGGGATGAGATTCAAAATACATCAGCTCCATCGATTTGGCTTGA 300
QY 97 LeuLeuLysProIleValLysGlnIleArgArgProLysHISValValValLysPheVal 116
Db 301 CCTATGAACCATCATTTAGGCBAATATACGAAGCCAAAGAAATGTGGTTCGCTAGCT 360
QY 117 ValLysPhePheProProAspHISThrGlnLeuGlnGlnGlnLeuThrArgTyTrLeuPhe 136
Db 361 GTAAATTTTTCACACTGATCTGTGTCACTACCAAGAAATATACAAATATCTGTTT 420
QY 137 AlaLeuGlnValLysGlnAspLeuAlaGlnGlyArgLeuThrCysAsnAspThrSerAla 156
Db 421 GCCCTGCACTTAAAGAGACCTGCTGGAAGACCGTTTGACTGTGTGACACACACAGCG 480
QY 157 AlaLeuLeuIleSerHISLysValGlnSerGlnIleGlyAspPheAspGlnAlaLeuAsp 176

Db 481 GCCCTTCCAGCTCCATCTCTCCAGTCCGAAATAGAGATTACAGAAACGTGGAC 540
Qy 177 ArgGluH1sLeuAlaLeuAsnLeuTyr1LeProGln1AspAlaLeuGluAsp1Le 196
Db 541 CGAGAGCACTCCAAAGTAAACGATATTGCTCGGCCAGCACTGCTTGAAGATA 600
Qy 197 ValGluPheH1sLeuAsnH1sLeuGln1PheProAlaGluSerPheGluLeu 216
Db 601 CTAGAAATCCATCAGACGCTGGGCCAGACACCTGCTGAGTCCGATTTCCAGGTGCTC 660
Qy 217 Glu1Lea1AspArgLeuGluMetTyrGly1LeArgLeuH1sProAlaLeuAspArgGlu 236
Db 661 GAAATGCTCGAAATGGAATTGAAATGACGGCATTTCACATTCAGTCTTGACGGAA 720
Qy 237 GlyThrLeu1LeuLeuAlaValAlaAsnThrGly1LeuVal1PheGlnGlyPheThr 256
Db 721 GGAACCAAGATTCAACTGGCAGATTCCCAATGGGTGTACTGTTCCAGNNNNNNNN 780
Qy 257 Lys1LeaAla1PheAsnThrPalalaValaLarglybLeuSerPheLysArgLysPhe 276
Db 781 NNN 840
Qy 277 Leu1LeuLysLeuArgProAspAlaAsnSerAlaTyrGlnAspThrLeuGluPheLeuMet 296
Db 841 NNN 900
Qy 297 AlaSerArgAspPheCysLeuSerPheThrLys1LeuVal1GluH1sLeuAlaPhePhe 316
Db 901 GGTAGTAGAGATGAATGTAAAGACTTCTGAAAGATTGTGTGAGATACACCTTTT 960
Qy 317 ArgLeuPheGluGluProLysPheProLysPheProValLeuPheSerArgGlySerSer 336
Db 961 AGACTTTTGGACCAACTAAGCCAAAGCAAAAGCCGTTCTTCAAGCCGGGCTCCCTCC 1020
Qy 337 PheArgSerArgLysArgThrGlnLysGlnValLeuAspTyrValLysGluGly1s 356
Db 1021 TTCAGATTACAGNN 1080
Qy 357 LysLeuVal1GlnPheGluArgLysH1sSerLys1Leu1s---Ser1AspSerLeu-- 374
Db 1081 NNN 1140
Qy 375 AlaSerGlnProThrGluLeuAsnSerGluValLeuGluGlnSerGlnSerThrSer 394
Db 1141 GCAGACTTACCAAAACAG-----AGC 1161
Qy 395 LeuThrPheArgLysGluGlyAlaGluSerProGlyGlnSerCysArgArg----- 411
Db 1162 ATCTCATTTCCCGAGGAGATTGAGACTCTGCTGCCCATCTTACGCCAATGCCCTTTAC 1221
Qy 412 -----GlyLysGluProLysValaSerAlaGlyGluProGly----- 423
Db 1222 TCGCTCTCTCCCTCACTCTGGTCCCTCTGGCTGCGCAGAGTTTAAAGACAGCAGCAGC 1281
Qy 424 -----SerH1sProSerProAlaProArgAspSerProAlaGlyAsnLysGlnAlaAsp 441
Db 1282 TCCCTCAAGATCCCGAGTTTCTCACTCAAGATCCAGCTGACAGAGAGGCGAGTGA 1341
Qy 442 GlyAlaAlaSerAlaProThrGlnGluGlnGluVal1Val1AspArgThrGlnGln 461
Db 1342 GCAGTGGCTTGAAGGCCCGAC-----ACACCATCG 1371
Qy 462 SerLysProGlnProPro-----GlnProSerThrGlySerLeuThrGlySerPro 478
Db 1372 GCCCGCCCTCGGGGCCCGCCGACCTGCTGCTGCGCTTTCACGAAAGAGTCT 1431
Qy 479 H1sLeuSerGluLeuSerValaAsnSerGlnGlyVala1AlaProAlaAsnVal1ThrLeu 498
Db 1432 CAGCTTCTCTCCCTCAGCGGAAAGAGCCCTGAGCTGAGCCCTGCA-----TTT 1482
Qy 499 SerProAsnLeuSerProAspThrLysGlnAlaSerProLeu1LeuSerProLeuAsn 518

Db 1483 CAGGTGCTTTGGGCCAGCTGAACAGGGCTCATGCCACTCTGAGCCCTGCTCAGT 1542
Qy 519 AspGlnAlaCysProArgThrAspAspGluAspGluGlyArgLysArgPheProThr 538
Db 1543 GATCTGGCGGAGCGCGGATGACTGCGAG--GAGCCACACACAAAGNNNNNNNNNN 1599
Qy 539 AspLysAlaTyrPhe1Lea1LysGlnValSerThrThrGluArgThrLysLeuAsp 558
Db 1600 NNN 1659
Qy 559 LeuGluVal1Leu1SerThrPheGlnSerThrValSerLysGluAspAlaMetProGlu 578
Db 1660 NNN 1719
Qy 579 AlaLeuLysSerLeu1LeuPheProAsnPheGluProLeuH1sLysPheH1sThrAspPhe 598
Db 1720 ACTGTAGAGCGCTCTTCTTCAACATTCATCCATCTATGAGTTTCACAGAGCTTC 1779
Qy 599 LeuLysGlu1LeuGlnArgLeuAlaLeuThrGluGlyArgSerAsnAlaGln1LeArg 618
Db 1780 CTGCGCGAGGTGGAGCAGAGCTGGCACTTGGGAAGGCCCTCCAAAGCCACACAAAA 1839
Qy 619 ---AspTyrGlnArg1Leu1AspValaMetLeuLysAsn1LeuGlnMetLysH1sLeu 637
Db 1840 GCAGATCATCAACGATCGGGGACATCCTGCTCAGAAACATGGCCAGTTAAAGAGATT 1899
Qy 638 AlaAlaH1sLeuThrLysH1sSerGluAlaLeuGlnAlaLeuGlnLysGly1LeuSer 657
Db 1900 ACCAGCTACTTCCAAAGCATGACGAGGTCTTAACAGAACTGGAAGGCTTACCAAGGC 1959
Qy 658 SerArgArgLeuGluAsnPheCysArgAspPheGluLeuGlnLysVal1CysTyrLeuPro 677
Db 1960 TGTAAAGATTGAGAGCGCTGACAGAGATTGAGCTGCGAAAGTCTGTACTGCT 2019
Qy 678 LeuAsnThrPheLeuLeuArgProLeuH1sArgLeuMetH1sTyrLysGlnValLeuGlu 697
Db 2020 CTCAACAGCTTCTGCTGAACCCATCGAGGGCTGCTGACATACCGCTGCTGCTGCGC 2079
Qy 698 ArgLeuCysLysH1sLeuProProSerH1sAlaAspPheArgAspCysArgAlaLeu 717
Db 2080 CGCCTATCGGACATTAACAGCCCGGCGACCATGACTGCTGACATGAGCGCTG 2139
Qy 718 AlaGlu1LeuThrGluMetVala1AlaGlnLeuH1sGlyThrMetLys1LeuMetGluAsnPhe 737
Db 2140 AAAGCATCACAGAGGTACACACACACTACAGACATTCATCCGGCTGAGAACTG 2199
Qy 738 GlnLysLeuH1sGluLeuLysLysAspLeu1LeuGly1LeaAsnLeuVala1ProGly 757
Db 2200 CAGAACTTAACGAGCTCAGCGGACCTGGTGGCATTAAGAACTCATTTGCTCTG 2259
Qy 758 ArgGluPhe1LeaGlyLeuGlySerLeuSerLysLeuSerGlyLysGlyLeuGlnArg 777
Db 2260 AGGAGTTCACTCGTGAAGGCTGCTTCAACAGCTCAACAAAGAGGCTTACAGAGG 2319
Qy 778 MetPhePheLeuPheAsnAspValaLeuLeuTyrThrSerArgLysLeuThrAlaSerAsn 797
Db 2320 ATGTTTTTCTGTTCTCAGATATGTTGCTGTACACAAAGAAAGAGTTGACGAGCAC 2379
Qy 798 GlnPheLysValaH1sGlyGlnLeuProLeuTyrGlyMetThr1LeuGlnSerGluAsp 817
Db 2380 CACTTCCGAGATCCGGGGCTCTTCCCTCCAAAGGCAATGCTGAGAAAGATGATAC 2439
Qy 818 GluThrGlyVal1ProH1sCysLeuThrLeuArgGlnArgGlnSer1Le1LeVala1a 837
Db 2440 GAGTGTCTGTTTCAACATGTTTCAACATTCACCGGCTCGAAACAAATGCTGTGCA 2499
Qy 838 AlaSerSerArgSerGluMetGluLysThrValaGlnAsp1LeuMetAla1AspLeu 857
Db 2500 GCCAGCATCTCGGCTGGAAGAAAGATGATGCTGACCTGAACCTCCGATCCAGCA 2559
Qy 858 AlaGluLysSerSerSerProAlaProGluPhe-----LeuAlaSerSerPro 873
Db 2560 GCCAAGAGTGGCGGTGACACGGCCCTGCACTGCGAGCGCGCAGCTGATGCACTGCTCC 2619

QY 874 ProAspAsnLysSerProAspGlnLaThrLaLaIAspGlnLysSerGluAspAsnLeu 893
|||
:::
QY 2620 -----AGATCCCCCAACAG--GTATCTGTGGACAGAGTCCAGAAATGAATCT 2670
|||
:::
QY 894 SerAlaSerArgThrSerLeuGluArgGlnLaProHisArgGluAsnThiMetValHis 913
|||
:::
Db 2671 CGGGGTGCCGAGAGCTCCCTGGAGGGGAGTGGCCAGCACCGGGGCCAACCAACAAATGAC 2730
|||
:::
QY 914 ValCysTrpHisValGluAsnThrSerValSerMetValAspPheSerIleAlaValGluAsn 933
|||
:::
Db 2731 GTGTGCTGTGTCCGGACACCAACGCTGTCCAGGGCAGACCAACGTGACGCTGTGAGANN 2790
|||
:::
QY 934 GlnLeuSerGluAsnLeuLeuArgLysPheLysAsnSerAsnGluTrpGlnLysLeuTrp 953
|||
:::
Db 2791 NNN 2850
|||
:::
QY 954 ValValPheThrAspAsnCysLeuPhePheGlyLysSerHisGlnAspAsnIleProLeu 973
|||
:::
Db 2851 NNN 2910
|||
:::
QY 974 AlaSerLeuProLeuLeuGluLysSerLeuThrIleProSerGluSerGluAsnIleGln 993
|||
:::
Db 2911 GCGACCTCCCGGCTGCTGGGCTACACGGGTGACATCCCAAGGAGAGGCGGATGGCTATAC 2970
|||
:::
QY 994 LysAspArgValPheLysLeuHisIlePheLysSerHisValLysArgPheArgAlaGluSer 1013
|||
:::
Db 2971 AAAGCATATGTTTCAAGTCCAGTTCAATCCCAAGTCACTTCTCCGGGCTGAGAGC 3030
|||
:::
QY 1014 GluLysTrpPheGlnLysArgTrpMetGluValIleArgSerAlaThrSerSerHisSerArg 1033
|||
:::
Db 3031 AAGTACATATTGAAAGGTGATGAGGTGATCCAGGGGTCACAGCTCCAGCCGGAGG 3090
|||
:::

RESULT 8	
AY415517	
LOCUS	AY415517
DEFINITION	AY15517 3168 bp DNA linear GSS 17-DEC-20
ACCESSION	Mus musculus FARP2 gene, VIRTUAL TRANSCRIPT, partial sequence,
VERSION	AY415517
KEYWORDS	genomic survey sequence.
ORGANISM	AY415517.1 GI:39771476
SOURCE	GSS.
REFERENCE	Mus musculus (house mouse)
AUTHORS	Mus musculus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidae; Muridae; Murinae; Mus.
AUTHORS	1 (bases 1 to 3168)
REFERENCE	Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J., Adams,M.D. and Cargill,M.
TITLE	Interferring nonneutral evolution from human-chimp-mouse orthologous gene trios
JOURNAL	Science 302 (5652), 1960-1963 (2003)
PUBMED	14671302
REFERENCE	2 (bases 1 to 3168)
AUTHORS	Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J., Adams,M.D. and Cargill,M.
TITLE	Direct Submission
JOURNAL	Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA
COMMENT	This sequence was made by sequencing genomic exons and ordering them based on alignment.
FEATURES	location/Qualifiers
SOURCE	1..3168
gene	/organism="Mus musculus" /mol_type="genomic DNA" /db_xref="taxon:10090" <1..3168 /gene="FARP2" /locus_tag="HCM5583"

ORIGIN			
Alignment Scores:			
Pred. No.:	1,92e-186	Length:	3168
Score:	2286.00	Matches:	466
Percent Similarity:	59.9%	Conservative:	167
Best Local Similarity:	44.1%	Mismatches:	376
Query Match:	41.8%	Indels:	48
Da:	14	Gaps:	9
US-09-555-342B-2 (1-1045) x AY415517 (1-3168)			
QY	1 MetGylgLuileglu-----glnaAgProthPrglYserArgleuGlyAla	16	
Db	1 ATGGGAGAGATGAAAGAAACATACAGAGCTTTTACCAACTTACGACACAGCTTGGGTGC	60	
QY	17 ProGluAanSerGlyIleSerThrleuGluArGlyGlnLysPrgProPrgProSer	36	
Db	61 CAGACCGGCATTGGAGTGAGTACCTCTGGAGCCAGACAGCTGTCCACCAGAGTGCAG	120	
QY	37 GlyLysLeuValSerIleYsrlleaglmeLteAspAspThrIngluAlaPheGluVal	56	
Db	121 GAGAAAGCACATCGCATCATCAGATGAAAGCTGCGAGACAGCACTGTAGATTATTGACATC	180	
QY	57 ProGlnaGAlaPrgGlyLysValleuLeuAspAlaValCysAenhiSleuAenLeuVal	76	
Db	181 GAGCCGAAGTGTGAAGCGGAGGTCTTACTGACTCAAGTGTGAAAGCATTTGAATCTCAT	240	
QY	77 GluGlyAspArgPheGlyLleuGlnPhePrgAspHisLysLysLethrValTrpLeuAsp	96	
Db	241 GAATGTACTACTTTGGTGTGGAGTTCAGAAATGTCCAGTCTTACTGAGATTGGCTTGA	300	
QY	97 LeuLeuLysProIleValLysGlnIleArGArGProLysHisValValLysPheVal	116	
Db	301 CCTATGAAACCCATCATATTAGCGAAAGTAAAGCAAAATAAGCGGTGCTGCGCTGGCA	360	
QY	117 ValLysPhePhePrgProPrgAspHisThrGlnLeuGlnGluLeuThrArgTyLLeuPhe	136	
Db	361 GTAAATATTTTCCGCCCTGACCCGTGTCAGCTGCAAGAAAGATACAAAGGTACTGT	420	
QY	137 AlaLeuGlnValLysGlnAspLeuAlaGlnGlyArgLeuThrCysAsnAspThrSerAla	156	
Db	421 GCTTTGCAACTAAAGAGACTCTCTGGAGGAACGGCTCACTGACCGGCCAACACTGCA	480	
QY	157 AlaLeuLeuIleSerHisIleValGlnSerGluIleGlyAspPheAspGluAlaLeuAsp	176	
Db	481 GCCCTTCATATCCCACTTCGTGCAGTCGGAATCGGAATGTATGATGAACACCTGGAT	540	
QY	177 ArgGluHisLysLeuAlaLysAsnLysTyrllePrgGlnGlnAspAlaLeuGluAspLysIle	196	
Db	541 CGAGAAACACCTCAAGCCCATGAATATCTGCCCAACAGAGAAATCTTACGAAAGATA	600	
QY	197 ValGluPheHisLysAsnHisIleGlyGlnThrProAlaGluSerAspPheGlnLeuLeu	216	
Db	601 CTAGACTTCATCAGAGGACACGGGCCAGACTCTCTGACAGATCAATTTCCAGGTGCTT	660	
QY	217 GluIleAlaArGArGleuGluMetTyrglyIleArGLeuHisPrgAlaLysAspArGlu	236	
Db	661 GAGATTGCAAGGAAGCTGGAATGTATGTGCATCAGGTTTCACATGGCTTCTGCACAGAA	720	
QY	237 GlyThrLysIleAsnLeuAlaValAlaAsnThrGlyIleLeuValPheGlnGlyPheThr	256	
Db	721 GGGACCAAGATTAAATCTAGAGTTTCTCACATGGGTGTCTGTGTCTCCAGNNNNNNNN	780	
QY	257 LysIleAsnAlaPheAsnTrpAlaLysValArgLysLysSerPheLysArgLysArgPhe	276	
Db	781 NNN	840	
QY	277 LeuIleLysLeuArGProAspAlaAsnSerAlaTyrglnAspThrLeuGlnPheLeuMet	296	
Db	841 NNN	900	
QY	297 AlaSerArGAspPheCysLysSerPheTrpLysIleCysValGluHisIleAlaPhePhe	316	

```

Db      901 GGTAGCAGATGATGATGAAGACTTGTGAAGATATGTGTGAGTACATACCTTTT    960
Qy      317 ArgLeuPheGluGluProLyPProLyPProValLeuPheSerArgLySerSer    336
Db      961 AGACTCTCTGACAGCCCTAGCCAAAGGCGTGTCTTCTCAAGCGAGGCTCCCTCC    1020
Qy      337 PheArgPheSerGlyArgThrGlnLeuGlnValLeuAspTyrValLeuGluGlyHis    356
Db      1021 TTCAGATACAGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN    1080
Qy      357 LysLysValGlnPheGluArgLysHisSerLysValHisSerLeuArgSerLeuLaser    376
Db      1081 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN    1119
Qy      377 GlnProThrGluLeuAsnSerGluValLeuGluGlnSerGlnGlnSerThrSerLeuThr    396
Db      1120 -----ACATCTCTTCATGCTCTGACTGTAGTCTGCTTAAACAGAGGCTCTTCACC    1173
Qy      397 PheGlyGluGlyAlaGlySerProGlyGlyGlnSerCysArgArgGlyLysGluProLys    416
Db      1174 GATGCTTGAGAACTTCTGCC-----TTC    1197
Qy      417 ValSerAlaGlyGluProGlySerHisProSerProAlaProArgArgSerProAlaGly    436
Db      1198 CTGCTTCAGCAAAATGCTCTCTTTATCCACCCCTAGTCTTCTTGTCCCTCTGCTGCT    1257
Qy      437 AsnLysGlnAlaAspGlyAlaAlaSerAlaProThrGluGluGluGlnValValLys    456
Db      1258 CTGCCCAATTGTAAGAGACAGACAGACAGCTCCCTTGATCCGACAGGCTCCGCTCAAG    1317
Qy      457 Asp-----ArgThrGlnGln    461
Db      1318 AGCAGACAGACAGAGAGACAGCGGACCATCATCTGACGGCCCAAGACACAGCTCG    1377
Qy      462 SerLys---ProGlnProPro-----GlnProSerThrGlySerLeuThrGlySerPro    478
Db      1378 GCCCATCTCCCTGAGACCCCGTGTGCTCGGCTGTCCAGGCTTTTCATGATAGTCT    1437
Qy      479 HisLeuSerGluLeuSerValAsnSerGlnGlyValAlaProAlaAsnValThrLeu    498
Db      1438 CAGCTTCTCCCTCCTCAAGACAGCCACTGAGCTGTGCTCT-----GAGCTT    1488
Qy      499 SerProAsnLeuSerProAspThrTyrLeuGlnLaserProLeuLaserProLeuLaser    518
Db      1489 CAGGCGCGCTGAGTACAGCTGACGAGGAGGTGATCCCTGTGCTCAAGCCGTGTCTGCT    1548
Qy      519 AspGlnAlaCysProArgThrAspAspGluAspGluGlyArgArgLysArgPheProThr    538
Db      1549 GGCCTGTGTACAGCCAGATGACCAACAGAGAGCAAGAAACACAGNNNNNNNNNNNN    1608
Qy      539 AspLysAlaTyrPheHisAlaLysGlnValSerThrThrGluArgThrTyrLeuLysAsp    558
Db      1609 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN    1668
Qy      559 LeuGluValIleThrSerTyrPheGlnSerThrValSerLysGluAspAlaMetProGlu    578
Db      1669 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN    1728
Qy      579 AlaLeuLysSerLeuLysPheProAsnPheGluProLeuHisLysPheHisThrAsnPhe    598
Db      1729 GCCCTGATGCGCTGCTTTCTCCAACTGTATCCAGTCTCAAGACTTCCACAGAGGCTTC    1788
Qy      599 LeuLysGluLysGluGlnArgLeuAlaLeuTyrGluGlyArgSerAsnAlaGlnLysArg    618
Db      1789 CTTCAACGAGTGAACAGAGGCTGGCACTGTGGAGGGGCTCCAGTGGCCCACTTAA    1848
Qy      619 ---AspTyrGlnArgGlyLysValMetLeuLysAsnLysGlnGlyMetLysHisLeu    637
Db      1849 GGTATATCCAGCAATCGGAGACATCTCTCCAGAGAACATGCGCACTTAAAGGAATTT    1908
Qy      638 AlaAlaHisLeuLysTyrLysHisSerGluAlaLeuGluAlaLeuGluAsnGlyLysSer    657

```

```

Db      1909 ACTAGTACTTCCAAAGACAGATGAGTCTCTTAAACAGAACTGTGAAGCCCAAAACAC    1968
Qy      658 SerArgArgLeuGluAsnPheCysArgAspPheGluLeuGlnLysValCysTyrLeuPro    677
Db      1969 TGTAAAGAGCTGGAGGACAGTCTACAAAGTTTGAAGCTCCAAAGGCTGTCTACTGCT    2028
Qy      678 LeuAsnThrPheLeuLeuArgProLeuHisArgLeuMetHisTyrLysGlnValLeuGlu    697
Db      2029 CTCAACACATTCCTGTGAAGCCGCTCCAGAGGCTAGTCCACTCCCTGCTGTGAGC    2088
Qy      698 ArgLeuCysLysHisHisProProSerHisAlaAspPheArgAspCysArgAlaAlaLeu    717
Db      2089 CGGCTGTGTCTCACTACTCTCTGAGCAGCGGACATATGCGGACTGTCATGAGGACACTG    2148
Qy      718 AlaGluIleThrGluMetValAlaGlnLeuHisGlyThrMetIleLysMetGluAsnPhe    737
Db      2149 AAGGCATCACAGAAAGTACACAGAGTCTCAGCAAAAGCTTACCGGCTGAAAACCTA    2208
Qy      738 GlnLysLeuHisGluLeuLysLysAspLeuIleGlyIleAspAsnLeuValProGly    757
Db      2209 CAGAAATTGACGAGCTACAGCCAGACCTGCTCGGTGATGAAGAACTCATTTGCTCTGG    2268
Qy      758 ArgGluPheIleArgLeuGlySerLeuSerLysLeuSerGlyLysGlyLeuGlnArg    777
Db      2269 AGGAGATTATCCGTGAGGCTGCTGCAACAGCTCACAAAGAAAGGCTCTCAACAGAG    2328
Qy      778 MetPhePheLeuPheAsnAspValLeuLeuTyrThrSerArgGlyLeuThrHisSerAsn    797
Db      2329 ATGTTTTTCTGTTCTCAGATATGTACTGTATCAAGCAAAAGTCTACAGAGCCAGT    2388
Qy      798 GlnPheLysValHisGlyGlnLeuProLeuTyrGlyMetThrIleGluGluSerGluAsp    817
Db      2389 CATTCGGAATCCGTGCTCTCTCCACTCCGTGCGATGCTGATGAAGAAAGTACAGAT    2448
Qy      818 GluTyrGlyValProHisCysLeuThrLeuArgGlyGlnArgGlnSerIleLeuValAla    837
Db      2449 GAATGCTGTGTTCTTCAATGCTTCACTATGACAGTCAAGAAACATATGTGTGAGCA    2508
Qy      838 AlaSerSerArgSerGluMetGluLysTyrValGluAspIleGlnMetAlaIleAspLeu    857
Db      2509 GCCAGCACTCGCTAGAAAGAAAGATGATGACAGACCTGAATGACAGCAATCAA---    2565
Qy      858 AlaGluLysSerSerSerProAlaProGluPheLeuAlaSerSerProProAspAsnLys    877
Db      2566 GCAGCCAAAGTATCGGTGACTACCCCACTGCTGTGGAGGCTCGGTATATCTGT    2625
Qy      878 SerProAspGluAlaThrAlaAlaAspGlnLysSerGluAspLeuSerAlaSerArg    897
Db      2626 ACCCTATGATCTCTGATGAAGTCTCTGTGAAAGATCAGAAAGTGTGAGAAACCGG    2685
Qy      898 ThrSerLeuGluArgGlnAlaProHisArgGlyAsnThrMetValHisValCysTyrHis    917
Db      2686 GGCCTCGTGGAGGGAACAGCCAGCACCGGCTCAATCAACAAATGATGTGTGTGATAC    2745
Qy      918 ArgAsnThrSerValSerMetValAspPheSerIleAlaValGluAsnGlnLeuSerCly    937
Db      2746 CGTATATCAAGTGTGTCCAGACGACACAGTCACTGTTGTGAGNNNNNNNNNNNNNN    2805
Qy      938 AsnLeuLeuArgLysPheLysAsnSerAsnGlyTyrGlnLysLeuTyrValPheThr    957
Db      2806 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN    2865
Qy      958 AsnPheCysLeuPhePheTyrLysSerHisGlnAspAsnHisPProLeuAlaSerLeuPro    977
Db      2866 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN    2925
Qy      978 LeuLeuGlyTyrSerLeuTyrIleProSerGluSerGluAsnIleGlnLysAspTyrVal    997
Db      2926 CTATTTGGTTACACGCTGAGGCTCTCCAGAGAGGCTACAGACTCACAAAGACTATGTC    2985
Qy      998 PheLysLeuHisPheLysSerHisValTyrTyrPheArgAlaGluSerGlyTyrThrPhe    1017
Db      2986 TTCAAGCTCCAAATCAATCAATCACTTCTTCCGGGCTGAGAGCAAGTACACATTT    3045

```



```

Db      1285 GCAGAGAGACAGTGCAGTGCCTGAGGCCCGAC----- 1323
Qy      456 LybAapAagThrGlnGlnSerIleuSerProGlnProP-----GlnProSerThrGly 472
Db      1324 -----ACACCATCGGCCAGCCCTCGGCGCCCGCAGCTCCAGCTCGTCCAGGC 1374
Qy      473 SerIleuThrGlySerProHlaIleuSerGluIleuSerValAanSerGlnGlyValAla 492
Db      1375 CTTTCCACGAAAGAGCTCTCAGCTTCTCCCTCCAGCCGGAAGAGCCCTCGAGTCTGAGC 1434
Qy      493 ProAlaenValThrIleuSerProAanIleuSerProAerThrLySGlnAlaSerProIleu 512
Db      1435 CCTGCA-----TTTCAGGTGCTTTGGCGCCAGCTGAACAGGGCTCATCCCACTC 1485
Qy      513 IleSerProIleuIleuAanIleuAanIleuAcyProAlaGthrAapAapGluAapGluAarg 532
Db      1486 CTGAGCCCTGCTCCAGTATGCTGCGGAGCCGGATGACTACGAG--GAGCCCAGA 1542
Qy      533 ArgLyAargPheProThrAapLyAalAthrPheIleAlaLySGluValSerThrGlu 552
Db      1543 CACAAANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1602
Qy      553 ArgThrTyrlleuLyAapLeuGluValIleThrSerTrpPheGlnSerThrValSerLyS 572
Db      1603 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1662
Qy      573 GluAapAlaMetProGluAlaIleuLySerIleuIlePheProAanPheGluProIleuHis 592
Db      1663 GAGAGCGCATGCTGCTGACTCTGATGACGCTCTCTCCAAATGATGATCCCATCTAT 1722
Qy      593 LyAerPheIleThrAanPheIleuLySGluIleGluGlnAargLeuAlaIleuTrpGluGlyAarg 612
Db      1723 GAGTTCCACAGAGCTTCTCGCGAGGTGAGAGAGCTGCACTTGGAAAGAGGCC 1782
Qy      613 SerAanIleuGlnIleAarg--AapTyrlnaGlnIleGlyAapValMetLeuLyAanIle 631
Db      1783 TCCAAAGCCCAACAAAGGAGCATCAAGAAATCGGGAGCATCTGTCGAGAAATG 1842
Qy      632 GlnIleuMetLyAanIleuAlaIleuIleuTrpLyAanIleuSerGluAlaIleuGluAla 651
Db      1843 CGCCAGTTAAAGGATTTACACAGCTCTCCAAAGACATGACGAGTCTTCAACAGAACT 1902
Qy      652 GluAanGlyIleuLySerSerAargAargLeuGluAanPheCysAargAapPheGluGln 671
Db      1903 GAAAGGCTACCAACGCTGTAAAGATTGAAGGAGCATGTACAAAGAGCTGAGCTGAG 1962
Qy      672 LyAanValCysTyrlleuProIleuAanThrPheIleuAargProIleuHisAargLeuMetHis 691
Db      1963 AAGGTCTGCTACTTGCCTCAACAGCTTCTGCTGAAGCCCATGACGAGCTGTCGAC 2022
Qy      692 TyrlleuGlnValIleuGluAargCysLyAanIleuIleuProPheSerHisAalAapPheAarg 711
Db      2023 TACCGCTGCTGCGCGCGCTTATGCGGACATTAACCCCGGAGCACTAGACTACCT 2082
Qy      712 AapCysAargAlaAlaIleuAlaGluIlePheGluMetValAlaGlnIleuHisGlyThrMet 731
Db      2083 GACTGCATGACGCGCTGAAGCCATCAAGAGGTGACCAACCACTACAGACATTTCTC 2142
Qy      732 IleuMetGluAanPheGlnIleuIleuHisGluIleuLyAapLeuIleGlyIleAap 751
Db      2143 ATCCGCTGCGAAGACCTGCAAGACTTAACGAGGCTGACGCGGAGCTGCTGGGTGAG 2202
Qy      752 AenIleuValIleProGlyAargGluPheIleAargLeuGlySerLeuSerLyAanSerGly 771
Db      2203 AACCTCATTTGCTCTGCGAGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2262
Qy      772 LySGlyLeuGlnIleAargMetPhePheIleuPheAanAapValIleuLeuTyThrSerAarg 791
Db      2263 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2322
Qy      792 GlyLeuThrAlaSerAanGlnPheLyAanIleGlyGlnIleuProIleuTyrlleuMetThr 811

```

```

Db      2223 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2382
Qy      812 IleGluIleuSerGluAapGluTrpGlyValProHlaCysLeuThrIleuAargGlyGlnAarg 831
Db      2383 GTGAAAGAAAGTATTAACAGATGGTCTGTTCACACCTGTTTACACATCTTACCGGCTCAG 2442
Qy      832 GlnSerIleIleValAlaIleAaSerAargSerGlnIleuGluTyrlleuValGluAapIle 851
Db      2443 AAACATCTGCTGCTGCGAGCCAGACATCGCTGGAAGAAAGATGGATCTGAGACTG 2502
Qy      852 GlnMetAlaIleAapLeuAlaGluLySerSerSerSerProAlaProGluPhe----- 868
Db      2503 AACTCGCGATCCAAAGCAAGATGCGCGGTGACANNCCCTCGACTCCGACGCGGC 2562
Qy      869 ---LeuAlaSerSerProProAapAanLySerSerProAapGlnAlaThrAlaAapGln 887
Db      2563 ACTCGTGCATCTCTCCCTCC-----AGATCCCAACGAG--GATCTCTGAGAGAG 2613
Qy      888 GluSerGluAapAapLeuSerAlaSerAargThrSerIleuGluAargGlnAlaProHlaAarg 907
Db      2614 GAGTCNAGAAATGATGCTCGGCGGTGTCGAGCTCCCTGGAGGGGACATGGCCAGACCCG 2673
Qy      908 GlyAanThrMetValIleValCysTrpHisAargAanThrSerValSerMetValAapPhe 927
Db      2674 GCCAACACACCAAAATGACGCTGCTGATCCGGAACACGAGCTGTCACAGGCGAGACAC 2733
Qy      928 SerIleAlaValGluAanGlnIleuSerGlyAanIleuAargLyAanPheAanSerAan 947
Db      2734 AGTCACACTGTGAGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2793
Qy      948 GlyTrpGlnIleuTrpValIlePheThrAanPheCysLeuPheTyrlleuSerHis 967
Db      2794 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2853
Qy      968 GlnAapAanHisProIleuAlaSerLeuProIleuGlyTyrlleuThrIlePheSer 987
Db      2854 NNNNATGATCTCCACATGCGCAGCTCCCGCTGCTGGGCTACANNNNNAGATCCCGCAG 2913
Qy      988 GluSerGluAanIleGlnIleuAapTyrlleuValPheLyAanIleHisPheLySerHisValTyr 1007
Db      2914 GAGGCGCATGCAATACCAAGACTATGTTTCAAGCTCCAGTCAATCCACAGCTCAC 2973
Qy      1008 TyrlleuAargAlaGluSerGluTyrlleuPheGluAargTrpMetGluValIleAargSerAla 1027
Db      2974 TTTCTCCGGGCGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 3033
Qy      1028 ThrSerSerAlaSerAarg 1033
Db      3034 NNCAGCTCAGCGCGGAGG 3051

RESULT 10
BX364606
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo
1 (bases 1 to 968)
L.I.W.B., Gruber,C., Jesse,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On May 5, 2003 this sequence version replaced gi:30370819.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Creteilux, CP 5706 - 91057 Evry cedex - FRANCE
Email: seqref@genoscope.cns.fr, web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime

```


Percent Similarity: 98.54
Best Local Similarity: 98.54
Query Match: 29.64
DB: 4
Gaps: 0

US-09-555-342B-2 (1-1045) x BX354921 (1-1000)

```
Oy 355 GLYHSELYSELYVALGHPHEGLUVRGLEYHISESERLYLEHISERTILEARXSERLEU 374
Db 2 GGACATTAAGAGAGTCAGTTTGAAAGAGACACACCAAGATTCTATTCGGAGCCTT 61
Oy 375 AIAEERGINPROTHRGULUENASERGLUVALLEUGLUGINSERGLINSERTHXSER 394
Db 62 GCTTCACAGCCTACAGACTGAATTCGGAAGTCTGGAGCACTCTCACAGACACCGACC 121
Oy 395 LEUTHRPHGILYGLUYALAGLUSERPROGLYGLINSERCYAARGXRGILYVSGLU 414
Db 122 CTTACATTTGGAGAGAGTGGCGGAATCTCCAGGGGGCCAGAGCTGCCGGCGAGAAAGGAA 181
Oy 415 PROLYSVALSERIAGLYGLUPROGLYSETHIAPROSERPROIAPROHARGXSERPRO 434
Db 182 CCGAAGTT-TCCGCCGGGAGCCGGGGTCCGACCCGAGCCCTGCCCGAGAGAGAGCCCC 240
Oy 435 AIAGLYASNYSGINLAASPGIYALAAIASERIALAPROTHRGULUGLUGLUVAL 454
Db 241 GCGGGTAAACAGACGAGCGGAGCGAGCCGCTCGGGCCCAAGAGAGAGAGAGGTC 300
Oy 455 VALIYASAPARTHRGINSERLYBPROGLINPROPROGLINPROSERTHRGYSELEU 474
Db 301 GTTAAAGATAGACCCAGACAGATMAACCTCAGCCCCCGACCAAGACAGAGCTCCCTG 360
Oy 475 THRGLYSERPROIIEUSERGLUENSERVALIENSERGLNGLYVALIAPROILA 494
Db 361 ACTGGCAGTCTCTCACCTTTCGAGCTGTCTGAACTGCGAGGGGAGTGCCCTGCCC 420
Oy 495 ASNVAILTHIRLEUSERPROASNLUSERPROBPTHRILYSGINLAISERPROLEUIESER 514
Db 421 AACTGACCTTGTCTCCCACTGAGCCCAACACAGCAGAGGCTCTCCCTGATACGC 480
Oy 515 PROLEULENAAENAPGINALCYAPROARGTHRAPAPRGIUAPRGILYARGXRGILY 534
Db 481 CCGCTGCTGAATGACACAGCTGCCGCCCGGACGAGCATGAGTAGAGGCGGAGGAG 540
Oy 535 ARGHEBPROTHRAPPLYVALATYTPHEHLEALYSGILVALISERTHTRGLUARGTHR 554
Db 541 AGATTCCCACTGATMAAGCGTACTCTACTAGCTAAGAGAGTCTTACCAACGAGCAGACA 600
Oy 555 TYRLEUYASAPLEUGLUVALIETHSERTRPHEGLINSERTHVALISERTLYSGLUASP 574
Db 601 TATCTGAAGGATCTCGAAGTTATCACTTCGTGGTTTCAGACACAGTGAGCAAGAGGAC 660
Oy 575 ALANETPROGLUALALEULYSESERLEUIEBHEPROENPHEGLUPROLEUIELYAPHE 594
Db 661 GCCATGCCGGAAGACGTAAGAAAGTCTCATATTCCGAAATTTTGAACCTTTGCACAAATTT 720
Oy 595 HISERHAPHELEULYSGILUIEGLUGINARGLEUALALEUTRPGILUGLYARGXSERASN 614
Db 721 CATACTAATTTTTCAGAGAAATTGAGAGACACTTGGCTGTGGAGAGCGCGCTCAAAAT 780
Oy 615 ALAAGINILEARXAPRYRGINARGILEGIYASPVAILMECLEULYASAMILEGINIYMET 634
Db 781 GCCCAATTCAGAGATTACCAAGAAATCGGCGATGCTCATGCTGAAGAACTTCAAGGCAATG 840
Oy 635 LYEHILEULALALAHILEUTRPLYEHISESERGLUALA-LEUGLUALALEUGLUANGI 654
Db 841 AAGCACTCGGCGGCTCACTCTGGAARCAACAGCAGCGCTTTGGAGGCGCSTGGAGAAATGG 900
Oy 654 YILLYSESERXARGTRGLEUGLUANPHECYARGASPHGILUGLUNLYSVVALCY 674
Db 901 AATCAAGAGCTCCGCGCGCTGAGAACTTCTGCAAGACATTTTAAGCTCGAGAAAGTGTG 960
Oy 674 STYRLEU 676
|||||
```

```
Db 961 TTACTYA 967
RESULT 12
LOCUS DV784833
DEFINITION Hw liver 37.050725 B09 Bos taurus CF-24-HW liver cDNA library Bos
taurus cDNA, mRNA sequence.
ACCESSION DV784833
VERSION DV784833.1 GI:82637709
KEYWORDS EST.
SOURCE Bos taurus (cattle)
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
Bovidae; Bovinae; Bos.
REFERENCE 1 (bases 1 to 1354)
Yoon,D.H., Lee,S.H., Park,E.W., Cho,Y.M., Lee,J.H., Kim,H.,
Kim,H.Y., Park,J.H. and Oh,S.J.
Gene Expression Profiling of the Bovine Liver, adipose, and
skeletal muscle
Unpublished (2005)
JOURNAL
COMMENT Contact: Dr. Du-Hak Yoon
National Livestock Research Institute, RDA
564 Omockhun-dong, Suwon, 441-350, Korea
Tel: 82 31 290 1593
Fax: 82 31 290 1792
Email: dhyoon@rda.go.kr.
FEATURES
source
1..1354
location:Qualifiers
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/sex="Four males mixed"
/cell_type="hepatocyte"
/dev_stage="24 months old"
/lab_host="XJLI-BlueRF" strain"
/clone_lib="Bos taurus CF-24-HW liver cDNA library"
/note="Organ: liver tissue, Vector: Uni-ZAPR; Site_1:
EcoRI; Site_2: Xho I"
ORIGIN
Alignment Scores:
Pred. No.: 6,31e-126 Length: 1354
Score: 1581.50 Matches: 309
Percent Similarity: 97.0% Conservative: 12
Best Local Similarity: 93.4% Mismatches: 7
Query Match: 28.9% Indels: 4
DB: 10 Gaps: 1
US-09-555-342B-2 (1-1045) x DV784833 (1-1354)
Oy 707 HISLAASPHARGASPCYARGALALALEUALAGLUILETHRGIMETVALIAGIN 726
Db 29 TACAGCAGCTTCAGAGGAGCTGCCAGCTTGCCAGACATCAACGAAATGAGGCCAG 88
Oy 727 LEUHSGLYTHMETILELYMETGLUASNPHGILNLYSEULYSGILULEULYBLYASP 746
Db 89 CTTACAGCTGAGAGATCAAGATGAGAAATTTCCAGAGCTGCAATGCAACAAAGAT 148
Oy 747 LEUIEGILYLEAPAELEULVALIAPROGLYARGILUPHEILEARGLEUGLYSERLEU 766
Db 149 TTGATTTGGCATTGACATCTTGATATCCAGAGAGGAAATTCATCCCTGGCGAGCTC 208
Oy 767 SERLYSEUSERGLYSGILYLEUGLUGINARGHEPHELEUPHEANAPVALLEU 786
Db 209 AGCAAGCTCTCCGGGAGAGGCGCTTCAACAGCAGCATGTTCTTCCGTTCACAGATGCTTG 268
Oy 787 LEUTYRTHXSERXARGILYLEUTHRALASERASGLINPHELYSVVALHISGLYGINLEUPRO 806
Db 269 CTGATCACAGCGCGGGGCTGACCGCTCGAATCAATTTAAAGTCCACGCGAGCTCCG 328
Oy 807 LEUTYRGLYMETTHRILEGILUGLUSERGLUASPGILUTRPGILYVALPROHISCYLEUTHR 826
|||||
```

Db 329 CTCTACGGCAGTACTGATGAGAAAGTGAAGAGGATGGGGCGTCCCCCACTGCTGACT 388
Qy 827 LeuArgGlyValIleArgGlnSerIleIleValAlaIleSerSerArgSerGluMetGluVal 846
Db 389 CTGGGGGGCCAGGCGCAGTCATGCTGTGTGGCCGCGCTTCTGCGCCGGAGATGGACAG 448
Qy 847 TTPValGluAspIleGlnMetAlaIleAspLeuAlaGluIlySerSerSerProAlaPro 866
Db 449 TGGGTGAGACATACATGATGGCATTGATGGCCGAGAAAGACAGCGGCCGCCGCC 508
Qy 867 GluPheLeuAlaSerSerProProAspAsnIlySerSerProAspGluAlaThrAlaAlaAsp 886
Db 509 GAGTTCCTGGCCAGCAGCCCCCGGACAAAGTCCCCCGAGAGGCGG---GGGGCGGAC 565
Qy 887 GlnGluSerGlnAspAspLeuSerSerAlaSerArgThrSerLeuGlnArgGlnAlaProHis 906
Db 566 CAGGAGTCGAGAGACAGCTCAGTGCCTCGCGACCTCGCTGAGAGCGCCAGGCCCGCAC 625
Qy 907 ArgGlyAsnThrMetValHisValCysTrpHisArgAsnThrSerValSerMetValAsp 926
Db 626 CGGGGCAACACATGATGACGTATGCTGGGACCGGCAACACAGCGTGCATGTGTGAC 685
Qy 927 PheSerIleAlaValGluAsnGlnLeuSerGlyAsnLeuLeuArgIlyPheIlyAsnSer 946
Db 686 TTCAGCGTGGCGGTGGAGAAATCAGCTGTCTGGGAACCTGTGAGGAATTCAAAACAGC 745
Qy 947 AsnGlyTrpGlnIlyLeuTrpValIlePheThrAsnPheCysLeuPhePheIlySer 966
Db 746 AACGGGTGGAGAGAGCTGTGGTGGTCTTACAGAACTTGTGCTTCTTCTACAAATCG 805
Qy 967 HisGlnAspAsnHisProLeuAlaSerLeuProLeuGlnIlyTrpSerLeuThrIlePro 986
Db 806 CACGAGACATACACCCCTTGGCAGCTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 865
Qy 987 SerGluSerGluMetIleGlnIlyAspTrpValPheIlySerIlyAspIlySerHisVal 1006
Db 866 TCGAGAGCCAGGACATCCACAAAGATTACGTGTTCAGCTGCAATTCAG-TCCACAGTG 924
Qy 1007 TyrTrpPheArgAlaGluSerGluTrpThrPheGluIlyArgTrp-MetGluValIle-ArgS 1026
Db 925 TACTACTCTCCGGGCCGAAAGCGAATACAGTTTGAAAGGTGGAATGAGTATCCGAA 984
Qy 1026 ealArhTrSerSerAlaSerArgProHis 1035
Db 985 ATGCAACCACTCGGCTCACGTCCAC 1013

RESULT 13
BM551766 1026 bp mRNA linear EST 20-FEB-2002
LOCUS AGENCOURT 6575294 NIH_MGC_98 Homo sapiens cDNA clone IMAGE:5479350
DEFINITION 5', mRNA sequence.
ACCESSION BM551766
VERSION BM551766.1 GI:18789071
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo
1 (bases 1 to 1026)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strauberg, Ph.D.
Email: cgarbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: LNCM2000 row: h column: 07

FEATURES High quality sequence stop: 675.
Location/Qualifiers
1..1026
source /organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5479350"
/issue_type="astrocytoma grade IV, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_1lb="NIH MGC 98"
/note="Organ: brain; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adapter: GGCAAGAG (G). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

ORIGIN
Alignment Scores:
Pred. No.: 7,348-123 Length: 1026
Score: 1544.00 Matches: 318
Percent Similarity: 93.0% Conservative: 2
Best Local Similarity: 92.4% Mismatches: 13
Query Match: 28.3% Indels: 11
DB: Gaps: 4
US-09-555-342B-2 (1-1045) x BM551766 (1-1026)

Qy 667 AspPheGluLeuGlnIlyValCysTrpLeuProLeuAsnThrPheLeuLeuArgProLeu 686
Db 3 GACTTGAGCTGCAGAGAGGTGTGTCTTACTACCGCTCAACACCTTCTCTGCGCCACTG 62
Qy 687 HisArgLeuMetHisTrpIlySerGlnValLeuGluArgLeuCysIlyHisIlyProSer 706
Db 63 CACGGCTCATGACATCAACAGAGGTCTGGAGCGGCTGTGCAACACACCGCGAGC 122
Qy 707 HisAlaAspPheArgAspCysArgAlaAlaLeuAlaGluIleThrGluMetValAlaGln 726
Db 123 CAGCGGACTTCAAGGAGCTGCGAGCGCGCTTGGCAGAGATCACGAGATGGTGGCAG 182
Qy 727 LeuHisGlyThrMetIleIlySerMetGluAsnPheGlnIlySerIlySerIlySerIly 746
Db 183 CTCGAGGTATGATGATCAAGATGAGAAATTCACAAAGCTGACCAACATCCAGAAAGAT 242
Qy 747 LeuIleGlyTrpAspAsnLeuValIleProGlyArgGluPheIleArgLeuGlySerLeu 766
Db 243 TTGATTGGCATTTGACATCTTGTGTTCGGGAGAGGAGTTCAATCCGTCTGGCAGCTC 302
Qy 767 SerIlyLeuSerGlyIlySerGlyLeuGlnIlyArgMetPhePheLeuPheAsnAspValLeu 786
Db 303 AGCAAGCTCTCGGGGAGAGGGGCTCCAGCAGCGCATTTCTTCTCTGTTCAACAGAGTCTG 362
Qy 787 LeuTrpThrSerArgGlyLeuThrAlaSerAsnGlnPheIlyValHisGlyGlnLeuPro 806
Db 363 CTATACACGAGCGCGGGGCTGACGGCTCCATCAAGATTAAAGTCCACGGGCGCTCCG 422
Qy 807 LeuTrpGlyMetThrIleGluGluSerGluIlyAspGluTrpGlyValIleProHisCysLeuThr 826
Db 423 CTCTAATGGCATGATGAG 482
Qy 827 LeuArgGlyIlyAspIleGlnMetAlaIleAspLeuAlaGluIlySerSerSerProAlaPro 846
Db 483 CTCGGGGCCAGGAGAGTCCATCATCTGTGGCCGACAGTTCTGTGTCAGATGAGAGAG 542
Qy 847 TTPValGluAspIleGlnMetAlaIleAspLeuAlaGluIlySerSerSerProAlaPro 866
Db 543 TGGGTGAGACATCAATGATGGCATTGATGGCCGAGAGAGAGAGAGAGAGAGAGAGAG 602
Qy 867 GluPheLeuAlaSerSerProProAspAsnIlySerSerProAspGluAlaThrAlaAlaAsp 886
Db 603 GAGTTCCTGGCCAGAGCCCCCTGACAAAGTCCCTGTATGAGCACAGCGGCTGAC 662

QY 887 GlnGluSerGluAspAspLeuSerAlaSerArgThrSerLeuGluArgAlaProHis 906
Db 663 CAGGATGAGAGAGTACCTGAGCCCTCGCGACATGCTGAGCCGACGCCCGGAC 722
QY 907 ArgGluAsnThrMetValHisValCysThrPHisArgAsnThrSerValSerMetValAsp 926
Db 723 CGCGGCAACACATGCTGCACTGCTGCTGCGCACCCCAACACAGAGCTCTCCATGCTGAC 782
QY 927 PheSerIleAlaValGluAsnGlnLeuSerGluAsnLeuLeuArgLysPheLysAsnSer 946
Db 783 TTCAGCATCGCAGTGAAGATCAGTGTCTGGAACCTGCTGAGGAGAAATTCAAAAACAGC 842
QY 947 Asn-GlyThrGlnLysLeuLeuTrp-Val-ValPheThrAsn---PheCysLeuPhePheTyr 964
Db 843 AACGGGGTGGCAAGAGCTGTGGGGTGGGTGTTCACAAAACCTTGCGCTGTCTTCTTAC 902
QY 965 LysSerHisGln---AspAsnHisProLeuAlaSerLeuProLeu-LeuGlyTyrSerIle 983
Db 903 AAATCACACCAAGGACATTCATCCCTTGCCTCCAGCCCTGCTGCTGCGGTTA---CT 959
QY 983 uThrIleProSerGluSerGlu-----AsnIleGlnLysAspTyrVal-PheLysL 1000
Db 960 CGGTTACACATCCCTCTGAAATCCCAAGACCTTCAGAAAGACTAGTGTCAAGC 1019
QY 1000 eu 1000
Db 1020 TT 1021

RESULT 14
LOCUS B0881649 912 bp mRNA linear EST 16-AUG-2002
DEFINITION AGNCOURT_8728249 NIH_MGC_47 Homo sapiens cDNA clone IMAGE:6339275
5', mRNA sequence.
ACCESSION B0881649
VERSION B0881649.1 GI:22273657
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 912)
AUTHORS NIH-MGC http://mgs.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgsrbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LMNL at:
http://image.llnl.gov
Plate: LICM2533 row: n column: 12
High quality sequence stop: 664.
Location/Qualifiers
1..912
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_image="6339275"
/tissue_type="neuroblastoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 47"
/note="Organ: brain; Vector: pOT7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dt priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adapter: GGCACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit

ORIGIN
Note: this is a NIH_MGC Library." (Stratagene) and Superscript II RT (Life Technologies).
Alignment Scores:
Pred. No.: 2,44e-120 Length: 912
Score: 1514.00 Matches: 296
Percent Similarity: 97.7% Conservative: 1
Best Local Similarity: 97.4% Mismatches: 7
Query Match: 27.7% Indels: 1
Gaps: 0
US-09-555-342b-2 (1-1045) x B0881649 (1-912)
QY 182 LysAsnLysTyrIleProGlnGlnAspAlaLeuGluAspLysIleValGluPheHis 201
Db 1 AAAAATAAATCAATACCTTCAGCAAGAGACACTTAAGACAAATCGGAAATTCACCAT 60
QY 202 AsnHisIleGlyGlnThrProAlaGluSerAspPheGlnLeuGluIleAlaArgArg 221
Db 61 AACCATTTGACAAACACACAGCAAGATTCAGATTTCCAGCTCTTAGAGATTGCCGTCG 120
QY 222 LeuGluMetTyrGlyIleArgLeuHisProAlaLysAspArgGluGlyThrLysLeuAsn 241
Db 121 CTAGAGATGTATGGAATCCGGTTGCACCCGCGCAAGGACAGGAGGACGACGATGCAT 180
QY 242 LeuAlaValAlaAsnThrGlyIleLeuValPheGlnGlyPheThrLysIleAsnAlaPhe 261
Db 181 CTGGCCGTGGCCAAACAGCGGAAATCTAGTGTTCAGGCTTTCACTAAGATCAATGCCCTTC 240
QY 262 AsnThrAlaLysValAlaArgLysLeuSerPheLysAspArgPheLeuIleLysLeuArg 281
Db 241 AACTGGGCGAAGGCGGAGAGCTAGCTTCAAGAGAAAGCCCTTCTCATCAAGCTCCGG 300
QY 282 ProAspAlaAsnSerAlaTyrGlnAspThrLeuGluPheLeuMetAlaSerArgAspPhe 301
Db 301 CCAATGCAATATATGAGTACAGAGATACCTTGGAATTCCTGATGGCCAGTCCGGAATTC 360
QY 302 CysLysSerPheThrPylIleCysValGluHisIleAlaPhePheArgLeuPheGluGlu 321
Db 361 TGCAGAGCTCTTCTGAAATCTGTGTGAACATCATGCTCTTTTAACTTTTGAAGAG 420
QY 322 ProLysProLysProLysProValLeuPheSerArgGlySerSerPheArgPheSerGly 341
Db 421 CCACAAACCAAGCCCAAGCCGCTCTTTAGCCGAGGATCATATTCGTTAGTGT 480
QY 342 ArgThrGlnLysGlnValLeuAspTyr-ValLysGluGlyHisLysLysValGlnPhe 361
Db 481 CGGACTCAGAAAGCAGGTTCTCGACTATGTTAAAGAGAGACATAAGAGGTGCAGTTT 540
QY 362 GluArgLysHisSerLysIleHisSerIleArgSerLeuAlaSerGlnProThrGluLeu 381
Db 541 GAAAGGAAGACACAGCAAGATTCATTCGAGGACCTTGTTCACAGCTACAGAACTG 600
QY 382 AsnSerGluValLeuGluGlnSerGlnGlnSerThrSerLeuThrPheGlyGluValAla 401
Db 601 AATTCGAAAGTGTGAGAGCTTCACAGAGACACAGCTTACATTGGAAGAGTGC 660
QY 402 GluSerProGlyGlyGlnSerCysArgArgGlyLysGluProLysValSerAlaGlyGlu 421
Db 661 GAATCTCCAGGGGGCCAGAGCTGCGCGGAGAAAGAACCGAAGGTTTCCGCCGGGAG 720
QY 422 ProGlySerHisProSerProAlaProArgArgSerProAlaGlyAsnLysGlnAlaAsp 441
Db 721 CCGGGGTGCAACCCGAGCCCTGCGCCAGAGAGAACCCCGGGGTAAACAAGCAGCGGAC 780
QY 442 GlnAlaLysSerAlaProThrGlnGluGlnGluValValLysAspArgThrGlnGln 461
Db 781 GAAACCCCTTTCGGCGCCGAGAGAGAAAGAGAGAGTCTTAAAGATGAGACCCACAG 840
QY 462 SerLysProGlnProProGlnProSerThrGlySerLeuThrGlySerProHisLeuSer 481
Db 841 AGTAAACCTCAGGCCCGGAGCCCAAGCAGAGGCTCCTGACTGGCACTC-CACCTTTTC 899

Qy 482 GluLeuSerVal 485
 Db 900 GAAGTGTCTGTGT 911

RESULT 15
 BU931960 904 bp mRNA linear EST 18-OCT-2002
 LOCUS BU931960
 DEFINITION AGENCOURT 10485886 NIH_MGC_109 Homo sapiens cDNA clone
 IMAGE:6672254 5', mRNA sequence.
 ACCESSION BU931960
 VERSION BU931960.1 GI:24120779
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homiidae; Homo.
 1 (bases 1 to 904)
 NIH-MGC http://mgs.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabs-remail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.lnl.gov
 Plate: LCM2952 row: h column: 14
 High quality sequence stop: 721.
 Location/Qualifiers
 1..904
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6672254"
 /tissue_type="teratocarcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_id="NIH_MGC_109"
 /note="Organ: ovary; Vector: pOTB7; Site 1: EcoRI; Site 2:
 XhoI; cDNA made by oligo-dT priming. Directionally cloned
 into EcoRI/XhoI sites using the following 5' adaptor:
 GGCAACGAC(G). Library constructed by Ling Hong in the
 laboratory of Gerald M. Rubin (University of California,
 Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
 Superscript II RT (Life Technologies). Note: this is a
 NIH_MGC Library."

ORIGIN

Alignment Scores:
 Pred. No.: 8.77e-119 Length: 904
 Score: 1496.00 Matches: 296
 Percent Similarity: 98.3% Conservative: 1
 Best Local Similarity: 98.0% Mismatches: 2
 Query Match: 27.4% Indels: 3
 DB: 3 Gaps: 0

US-09-555-342B-2 (1-1045) x BU931960 (1-904)

Qy 323 LysProLysProLysProValLeuPheSerArgLysSerPheArgPheSerGlyArg 342
 Db 1 AAACCAAGCCCAAGCCCGTCTTTCAGCCGGGGTCATCATTTCCGTTCAAGTGTCCG 60

Qy 343 ThrGlnLysGlnValLeuAspTyrValLysGlnGlyGlyHisLysLysValGlnPheGln 362
 Db 61 ACTCAGAGCAGGTTCTCGACTTAAAGAGAGAGACATAAGAGTGCAGTTGAA 120

Qy 363 ArgLysHisSerLysLysLeuSerLysLeuSerLeuLysSerGlnProThrGlnLeuAsn 382
 Db 121 AGGAAGACACGACAGATTTCATTCGCGAGCCTTGCTTCACAGCCTACAGAACTGAAT 180

Qy 383 SerGluValLeuGlnLysSerGlnLysSerThrSerLeuThrPheGlnGlyGlnValGln 402
 Db 181 TCGGAAGTGTGAGCAGCTCTCAGCAGACGACACGACCTTACATTTGAGAAAGGTGCCGA 240

Qy 403 SerProGlnGlyGlnSerCysArgArgGlyLysGlnProLysValSerLysGlnPro 422
 Db 241 TCTCCAGGGGGCCAGAGCTGCCCGCAGAGAAAGAAACCGAAGGTTTCCCGGGGAGCCG 300

Qy 423 GlysSerHisProSerProLysProLysProLysProLysProLysProLysProLys 442
 Db 301 GGGTCCGACCCGAGCCCTGCGCAGAGAGAGCCCGGGGTAAACAGCGGCGACGGA 360

Qy 443 AlaAlaSerLysLysProThrGlnLysGlnLysGlnValValLysAspArgThrGlnLys 462
 Db 361 GCCGCTCCGGCCGCCAGAGGAAAGAGAGAGTGTGTTAAGATAGACACCGACAGAGT 420

Qy 463 LysProGlnProProGlnProSerThrGlySerLeuThrGlySerProHisLeuSerGln 482
 Db 421 AAACCTCAGCCCCCGCAGCCAGCAGCAGCAGCTCCCTGACTGGCAGTCTCACTTCCGAG 480

Qy 483 LeuSerValAsnSerGlnGlyValAlaProAlaAsnValThrLeuSerProAsnLeu 502
 Db 481 CTGTCTGTGAATCCGAGGGGGAGTGGCCCTGCCAAGCTGACTTGTCTCCCAACTG 540

Qy 503 SerProAspThrLysGlnLysSerProLeuLysSerProLeuLysAsnAspGlnAlaCys 522
 Db 541 AGCCCGACACCAAGCAGGCTCTCCCTTGATCAGCCGCTGTGATGACACGAGCCTGC 600

Qy 523 ProArgThrAspAspGlnAspGlnGlyArgArgLysArgPheProThrAspLysAlaTyr 542
 Db 601 CCCCGACGACGATGAGATGAGGGCCGAGAGAAAGATTTCCCACTGATTAAGCGTAC 660

Qy 543 PheIleAlaLysGlnValSerThrThrGlnArgThrTyrLeuLysAspLeuGlnValIle 562
 Db 661 TTCTATGCTTAAAGAAAGTGTCTACACCGACGACCAATCTGAAAGATCTCGAAGTTATC 720

Qy 563 ThrSerTyrPheGlnSerThrValSerLysGlnAspAlaLeuProGlnAlaLeuLysSer 582
 Db 721 ACTTCGGGTTTCAGAGCAGACAGTGAACAAAGAGAGCGCATGCCGAGACATGAAGAT 780

Qy 583 LeuIlePheProAsnPheGlnProLeuHisLysPheHisThrAsnPheLeuLysGlnIle 602
 Db 781 CTGATATTTCCGAAATTTTGAACTTTGCACAAATTTCTACTAATTTCTCAAGGAATTT 840

Qy 603 GlnGlnArg-LeuAlaLeu-TyrGlnGly-ArgSerAsnLysGlnLysLeuArgAspTyrGln 621
 Db 841 GAAGAACGACCTCCCTGGTGGAAAGCCCGCTCAAAATGCCAAATCAGAGATTACAA 900

Search completed: May 20, 2006, 08:51:24
 Job time : 11281 secs

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - nucleic search, using frame_p2n model

Run on: May 20, 2006, 02:18:42 ; Search time 449 Seconds

(without alignments)
6532.212 Million cell updates/sec

Title: US-09-555-342B-2

Perfect score: 5463
Sequence: 1 MSIEEDRPTGSRGLGAPENS.....SATSSASRPHVLSHKESLVY 1045

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 1403666 seqs, 935554401 residues

Total number of hits satisfying chosen parameters: 2807332

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=framed_p2n.model -DEV=x1h
-Q=/abs/ABSSMB.spool/US09555342/runat_19052006_121500_13276/app_query.fasta-1
-DB=Issued Patents NA -OPMT=fastcap -SUFFIX=p2n.rn1 -MISMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdt
-LIST=45 -DOCALL=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -NOM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-HOST=abssmb2h -USRR=US09555342_@CGN_1_1_204_@runat_19052006_121500_13276
-NCPU=6 -ICPU=3 -NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -MAIN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

- Issued Patents NA.*
1: /EMC_Celerra_SIDS3/prodata/2/ina/1.COMB.seq.*
2: /EMC_Celerra_SIDS3/prodata/2/ina/5.COMB.seq.*
3: /EMC_Celerra_SIDS3/prodata/2/ina/6.COMB.seq.*
4: /EMC_Celerra_SIDS3/prodata/2/ina/6.COMB.seq.*
5: /EMC_Celerra_SIDS3/prodata/2/ina/7.COMB.seq.*
6: /EMC_Celerra_SIDS3/prodata/2/ina/8.COMB.seq.*
7: /EMC_Celerra_SIDS3/prodata/2/ina/9.COMB.seq.*
8: /EMC_Celerra_SIDS3/prodata/2/ina/10.COMB.seq.*
9: /EMC_Celerra_SIDS3/prodata/2/ina/11.COMB.seq.*
10: /EMC_Celerra_SIDS3/prodata/2/ina/12.COMB.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	724	13.3	4292	US-09-949-016-1488 Sequence 1488, App
2	724	13.3	4336	US-09-949-016-208 Sequence 208, App
3	713	13.1	2853	US-09-949-016-1933 Sequence 1933, App
4	712.5	13.0	6263	US-09-664-958-5 Sequence 5, Appl1
5	712.5	13.0	6268	US-09-566-921-57 Sequence 57, Appl1
6	626	11.5	3398	US-09-799-451-201 Sequence 201, App
7	563	10.3	2872	US-09-906-779-3 Sequence 3, Appl1
8	557.5	10.2	2493	US-10-104-047-41 Sequence 41, Appl1

9	541	9.9	3984	US-09-848-294-1	Sequence 1, Appl1
10	508	9.3	2156	US-09-949-016-3593	Sequence 3593, App
11	503.5	9.2	576	US-09-270-767-1995	Sequence 1995, App
12	503.5	9.2	576	US-09-270-767-17277	Sequence 17277, App
13	497.5	9.1	2931	US-10-094-749-1401	Sequence 1401, App
14	443	8.1	4272	US-09-949-016-5747	Sequence 5747, App
15	424.5	7.8	4092	US-09-566-921-115	Sequence 115, App
16	405.5	7.4	2170	US-09-854-133-728	Sequence 728, App
17	403.5	7.4	3438	US-10-164-595-29	Sequence 29, Appl
18	399.5	7.3	1979	US-09-949-016-4612	Sequence 4612, App
19	396.5	7.3	4080	US-08-446-345-35	Sequence 35, Appl
20	396.5	7.3	4081	US-09-949-016-1292	Sequence 1292, App
21	382	7.0	3044	US-09-814-915A-84	Sequence 84, Appl
22	382	7.0	3044	US-09-880-107-3718	Sequence 3718, App
23	359.5	6.6	2080	US-08-179-738-9	Sequence 9, Appl1
24	359.5	6.6	2080	US-08-628-145-9	Sequence 9, Appl1
25	359.5	6.6	2257	US-08-171-718-15	Sequence 15, Appl
26	359.5	6.6	2257	US-08-478-087-15	Sequence 15, Appl
27	343	6.3	1862	US-08-179-738-4	Sequence 4, Appl1
28	343	6.3	1862	US-08-628-145-4	Sequence 4, Appl1
29	338.5	6.2	1817	US-08-179-738-1	Sequence 1, Appl1
30	338.5	6.2	1817	US-08-628-145-1	Sequence 1, Appl1
31	338.5	6.2	1833	US-08-179-738-6	Sequence 6, Appl1
32	338.5	6.2	1833	US-08-628-145-6	Sequence 6, Appl1
33	337	6.2	310	US-09-621-976-8193	Sequence 8193, App
34	332.5	6.1	2723	US-10-104-047-1219	Sequence 1219, App
35	331	6.1	573	US-09-270-767-165	Sequence 165, App
36	331	6.1	573	US-09-270-767-15447	Sequence 15447, App
37	322	5.9	56516	US-08-996-306-1	Sequence 1, Appl1
38	322	5.9	56516	US-09-338-907-1	Sequence 1, Appl1
39	322	5.9	56516	US-09-218-207-1	Sequence 1, Appl1
40	322	5.9	56520	US-09-338-907-179	Sequence 179, App
41	322	5.9	56520	US-09-218-207-179	Sequence 179, App
42	321	5.9	1626	US-09-620-312D-1033	Sequence 1033, App
43	321	5.9	1699	US-09-913-171A-26	Sequence 26, Appl
44	317.5	5.8	449	US-09-270-767-1542	Sequence 1542, App
45	317.5	5.8	449	US-09-270-767-16824	Sequence 16824, App

ALIGNMENTS

RESULT 1
US-09-949-016-1488
Sequence 1488, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: C1001307
CURRENT APPLICATION NUMBER: US/09/949, 016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1488
LENGTH: 4292
TYPE: DNA
ORGANISM: Human
US-09-949-016-1488
Alignment Scores:
Pred. No.: 4,58e-58
Score: 724.00
Percent Similarity: 44.0%
Best Local Similarity: 28.8%
Query Match: 13.3%
Length: 4292
Matches: 221
Conservative: 117
Mismatch: 287
Indels: 143
Gaps: 24


```

Oy      664 eCyatgaaapheglnleuglnlyvalcyvryrleuproleaasnthrphenleuat 664
Db      2698 -----CCATTGTCCAAACTGAGACCAAA 2720
Oy      684 gPro-----LeuHiaRgkueethiaryrlysglnvalleuglnaryleucy 700
Db      2721 ACCATTCACATATAGAGCTCTCCACAGATTGATGAGCGGGGCTGTGTATTTGGGACGTTA 2780
Oy      700 eLyshHiaProProSerHis 707
Db      2781 CTGACCGCACAAACCATCACAT 2802

RESULT 2
US-09-949-016-208
/ Sequence 208, Application US/09949016
/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
/ FILE REFERENCE: CL001307
/ CURRENT APPLICATION NUMBER: US/09/949,016
/ PRIOR FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 208
/ LENGTH: 4336
/ TYPE: DNA
/ ORGANISM: Human
US-09-949-016-208

Alignment Scores:
Pred. No.: 4,67e-58 Length: 4336
Score: 724.00 Matches: 221
Percent Similarity: 44.0% Conservative: 117
Best Local Similarity: 28.8% Mismatches: 287
Query Match: 13.3% Indels: 143
Db: 3 Gaps: 24

US-09-555-342B-2 (1-1045) x US-09-949-016-208 (1-4336)
Oy      34 ThrProSerGlyLyseuValserIleuValIleuGlnMetIleuAspAphThrglnIuaLa 53
Db      678 ACCAAGAAAGACCAAACTGTCAGTGTAAGAGACCTCTTAGATGGCACCGAAATACAGC 737
Oy      54 phegluValProGlnaRgHiaProGlyLyseuValleuAspAlaValCyaaHniIleu 73
Db      738 TGTGACCTGGAGAAACATGCCAAGGACCAAGTGTTATTGACAAAGTGTAACACCTC 797
Oy      74 AsnIleuValGluGlyAspTyrrPheGlyLeuGluPheProAspHiaValyValIethrVal 93
Db      798 AATCTCTTGGAGAAAGACTACTTGTGACTTTGTTTCAGGAAAGCCCTGACGAGAAAAC 857
Oy      94 TrpIleuAspLeuLeuLyseuProIleuValLyseuGlnIleArgArgProLyshHiaValVal 113
Db      858 TGGTTAAGATCCGCTTAAGAAATTAAGAAACAACTGAGAAACCTTCATGCGCTATTCCACT 917
Oy      114 LyseuPheValValLyseuPheProProAspHiaethrGlnIleuGlnGluLeuTharG 133
Db      918 ---TTTAATGTGAAGTTTATCTCTCTGTGATCCCTTCATATGACTGAGATATACACAGA 974
Oy      134 TyLeuPheAlaIleuGlnValLyseuGlnAspAlaGlnGlyArgIleuThrCyaaAsp 153
Db      975 TACTTCTTGTGCTTCAGCTCCGGCAGGCAATTTGCTTGGCGGCTGCTGCTGCTTTT 1034
Oy      154 ThrSerAlaAlaIleuLeuIleSerHiaIleValGlnSerGluIleGlyAspPheAspGlu 173

```

Db	1035	GTGACTACGCTCTTCGGGATCTCTACCCCTGACGGCTGAATCTGGATATGAC--	1091
Qy	174	AlaLeuAspArgLysHis-----LeuAlaLysAsnLysTyrIlePro---Gln	188
Db	1092	-----CCAGAAGAAACATGGACAGCATGCACCTCACTGAATTCACGTTGCCCTACATCAG	1145
Qy	189	GlnAspAlaLeuGlnAspLysIleValGlnPheHisAsnHisIleGlyGlnThrPro	208
Db	1146	ACTAAGAGAGCTGGAAAGAGAGGTGGCAGAGCTGCACAAACCCACAGGGGCTTATCCGCA	1205
Qy	209	AlaGlnSerAspPheGlnLeuLeuGlnIleLeuAlaArgArgLeuGlnMetTyrGlyIleArg	228
Db	1206	GCACAAAGCTGATCCCACTTGTGAAATGCAAAGAGCTTCCATGTAAGTGGTGGAC	1265
Qy	229	LeuHisProAlaLysAspArgGlnGlyThrLysIleAsnLeuAlaValAlaAsnThrGly	248
Db	1286	CTACATCATGCCAAGACACTCAAGAGGTGGACATCAAGCTGGGGCGTGTGCTAAATGA	1325
Qy	249	IleLeuValPheGlnGlyPheThrLysIleAsnAlaPheAsnTyrAlaLysValArgLys	268
Db	1326	CTTCTCATTTCAAAAGACAGACTCGAATCAATCGTTTGGCTGGCGAAATCTTAA	1385
Qy	269	LeuSerPheLysArgLysArgPheLeuIleLysLeuArgProAspAlaAsnSerAlaTyr	288
Db	1386	ATTTCCTTAACCGACAGTAACTTCTTACATTAAATGCACCGGACAGAGCTGGAACAGTTT	1445
Qy	289	GlnAspThrLeuGlnPheLeuMetAlaSerArgAspPheCysLysSerPheThrLysIle	308
Db	1446	GAGAGTACCTTGATTCATTAACCTGCACAAACCCGGGACGGAAAGACTATGAAAGTG	1505
Qy	309	CysValGlnHisIleAlaPhePheArgLeuPheGlnGluProLysProLysPro	328
Db	1506	TGGCGGAGGACTCATCTTCTTCTACAGGGTTGT--TCTCCAGAGCAGCACCAAAAGCC	1562
Qy	329	ValLeuPheSerArgLysSerPheArgPheSerGlyArgThrGlnLysGlnValLeu	348
Db	1563	AAGTTCCTGCAGCTTGGGGTCCAAATTTTCGCTATAGTGGCCGACCAAGCACACACCGC	1622
Qy	349	AspTyrValLysGlnGlyGlyHisIleLysLysValGlnPheGlnArgLysHisSerLysIle	368
Db	1623	CAGGCCAGCACCTCATATGATGATGGCCAGCACACACACTTGTAGGCCACTTGTATGAACGG	1682
Qy	369	HisSerIleArgSerLeuAlaSerGlnProThrGlnLeuAsnSerGluValLeuGlnGln	388
Db	1683	GTCTCC--AGGAGTCTATGATGAGCTCCGATT-----GGGTCTATGGACCAA	1727
Qy	389	SerGlnGlnSerThrSerLeuThrPheGlyGlnGlyAlaGlnSerProGlyGlyGlnSer	408
Db	1728	AGTCTTATGAAG-----GATTTTCCGTGGCGCTGTGGG	1760
Qy	409	CysArgArgGlyLysGluProLysValSerAlaGlyLysProGlySerHisProSerPro	428
Db	1761	-----GAAATTTTCAGCCATATGAGCACTTGACATT--GTCCAGCAATT	1796
Qy	429	AlaProArgArgSerProAlaGlyAsnLysGlnIleAspGlyAlaIleSerAlaPro---	447
Db	1797	GCCTGTGACAAAGTGGGACGGCAGAGAGGAAATGAGAAGCCCACTAAAGCCCAAT	1856
Qy	447	-----	447
Db	1857	TTGCAGCTCATTTGAAGAAAGAAATTTCTTGAGATAGAGGAGATTAATTATTATGTC	1916
Qy	448	-----ThrGlnGlnGlnGlnGlnValValValLysAsp	457
Db	1917	AGACATAGCAATTTAATGTTGGAGAACTGATGAGACAGAGGACATATCTGAACAT	1976
Qy	458	ArgThr-----GlnGlnSerLysProGlnProPro---	467
Db	1977	CAGGCTAGCATTAATGAATCAACGGCGCAATTTTATGGAATTCACACTGAGCCGGCCCT	2036
Qy	468	-----GlnProSerThrGlySerLeuThrGlySerProHis	479

```

Db      2037 AATGAATGGGAAAAACGACCTATCACACTCTGCTCCCTACAGACACAGGAGATTACAT 2096
Qy      480 -----LeuSerGluLeuSerValAsnSerGlnGlyValAlaProAlaAsn 495
Db      2097 GAGACTCTGAATATATGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2156
Qy      496 ValThr-----LeuSerProAsnLeuSerPro--AspThrLys 507
Db      2157 ATCACAGAAAGAAATGAATGTGTAAGATATACCTCGAGGTGGTCTCGGAGGAGATTGGT 2216
Qy      508 GlnAlaSerProLeuLe-----SerProLeuLeuAsnSerGlnAlaCysPro 523
Db      2217 AAGGTGAGGCTGTGACACAAAGAGCTCCCTGCTCTTCTTGAGACACAGCAGC 2276
Qy      524 ArgThrAspArgGluAspArgGluArgArgLysArgPheProThrAspLysAlaTyrPhe 543
Db      2277 AGCAGTGAAGTGAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2327
Qy      544 IleAlaLysGluValSerThrThrGluArgThrTyrLeuLysAspLeuGluValIleThr 563
Db      2328 GTGACCGAGGACACCATCAGAGAGAGAGAGAGAGAGATAGAGAGAGGTGAGAGAAACC 2387
Qy      564 SerTrpPheGlnSerThrValSerLysGluAspAlaMetProGluAla----- 579
Db      2388 CGCCCGCAGCCAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2447
Qy      580 -----LeuLysSerLeuLeu--PheProAs 587
Db      2448 CAAGCAGGTCCAGTGAATCACAGTGAAGAAACAGTATCCAGAGAAATGATAGTCCCAA 2507
Qy      587 nPheGluProLeuHisLysPheHisThrAsnPheLeuLysGluIleGlnGlnArgLeuAl 607
Db      2508 AAGATACCCGAGAGAGAGAGAGTGAACAGAGGCGCTCTTAAG-----AA 2552
Qy      607 AlaLeuTrpGluGlyArgSerSerAsnAlaGlnIleArgAspTyr--GlnArgIleGlyAspVa 626
Db      2553 GACATGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2612
Qy      626 IMetLeuLysAsnIleGlnGlyMetLysHisLeuAlaAlaHisLeuTrpLysHisSerG 646
Db      2613 GACATTGATTTTGGCCAAATATATTTGTTTTCAGAGCCACAGTGGTAAAAACAGAG 2672
Qy      646 ValAlaLeuGluAlaLeu-----GlnAsnGlyIleLysSerSerArgArgLeuGlnAsn 664
Db      2673 ATGTAACAATTTCTGATGCTCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2727
Qy      664 EcysAspArgPheGluLeuGlnLysValCysTyrLeuProLeuAsnThrPheLeuLeuAr 684
Db      2728 -----CAspTrpCCAACTGAGCCAA 2750
Qy      684 GPro-----LeuHisArgLeuMetHisTyrLysGlnValLeuGlnLysArgLeuCy 700
Db      2751 ACCATCAACATATAGTCTCCACAGATTGATGGGGGCTGTGTGATTTGGGGCAGCTTA 2810
Qy      700 ValHisHisLeuProProSerHis 707
Db      2811 CTGACCGCACAACCATCAT 2832

RESULT 3
US-09-949-016-1933
; Sequence 1933, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTNER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03

```

```

; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1933
; LENGTH: 2853
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-1933

Alignment Scores:
Pred. No.: 2,52e-57 Length: 2853
Score: 713.00 Matches: 190
Percent Similarity: 46.1% Conservative: 90
Best Local Similarity: 31.3% Mismatches: 197
Query Match: 13.1% Indels: 130
DB: 3 Gaps: 17

US-09-555-342B-2 (1-1045) x US-09-949-016-1933 (1-2853)
Qy      16 AlaProGluAsnSerGlyLysSerThrLeuGluArgGlyGlnLysProProThrPro 35
Db      719 AGCCCTCAATCAAAAGCAGAAACAGAAATTAAGCTTCCCAAAA-----CCAATCAGA 772
Qy      36 SerGlyLysLeuValSerIleLysIleGlnMetLeuAspAspThrGlnAlaPheGlu 55
Db      773 AAACACAGAGAACATGCATCGAAGGTTTCTTGTGGATGACACAGTTTATGAAATGTGTT 832
Qy      56 ValProGlnArgAlaProGlyLysValLeuLeuAspAlaValCysAsnHisLeuAsnLeu 75
Db      833 GTGGAGAAACATGCTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 892
Qy      76 ValGluGlyAspTyrPheGlyLeuGluPheProAspHisLysLysIleThrValTrpLeu 95
Db      893 TTGGAGAGAGACTATTTTGTCTAGCCATTGGGATTAAGCAACCTTAAAGCATGGCTG 952
Qy      96 AspLeuLeuLysProIleValLysGlnIleArgArg--ProLysHisValValValLys 114
Db      953 GATTCGCCCAAGAAATAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1006
Qy      115 PheValValLysPhePheProProAsnHisThrGlnLeuGlnGluLeuThrArgTyr 134
Db      1007 TTTAATGTAAAGTTTATACCTGACCCAGACAGTGAACAGAAACATTAACAAAGATAT 1066
Qy      135 LeuPheAlaLeuGlnValLysGlnAspLeuAlaGlnGlyArgLeuThrCysAsnAspThr 154
Db      1067 TATTTATGCTTTCAGGTTGGCAGACAGATGTCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1126
Qy      155 SerAlaAlaLeuLeuLysSerHisLysIleValGlnSerGlnIleGlyAspPheAspGluAla 174
Db      1127 ACCTTAGCATTTATTAAGTTCTTAACACATCCAGTCCGAACTGGAGAGACTAAGCCACAGA 1186
Qy      175 LeuAspArg--GlnHisLeuAlaLysAsnLysTyrIleProGlnGln--AspAlaLeu 192
Db      1187 CTCATGCGGTGATATGATGATGATTTTAACTGGCCCAATGAGCAAGAGAACTT 1246
Qy      193 GlnAspLysIleValGlnPheHisLysAsnHisIleGlyGlnThrProAlaGlnSerAsp 212
Db      1247 GAAAGAGAGAGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1306
Qy      213 PheGlnLeuLeuGlnIleAlaArgArgLeuGlnMetTyrGlyTyrLeuGlnHisProAla 232
Db      1307 TTGGAGTTTCTTGAGAGATGCCAAAAAGTGTCTATGATATGAGAGTTGATCTTCAATAAGCA 1366
Qy      233 LysAspArgGluGlyThrLysIleAsnLeuAlaValAlaAsnThrGlyLysLeuValPhe 252
Db      1367 AAGGACTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1426
Qy      253 GlnGlyPheThrLysIleAsnAlaPheAsnTrpAlaLysValArgLysLeuSerPheLys 272
Db      1427 AAAAGTAAAGTGAAGATTAACCGCTTCCCTTGGCCCAAGTGTGAAGATTTCTTATAAA 1486
Qy      273 ArgLysArgPheLeuIleLysLeuArgProAspAlaAsnSerAlaTyrGlnAspThrLeu 292

```

```

Db 1487 CGTAGAGCTTTTCATCATGATTCGGCTCGAGAGCAAGCAGTGTGAAGTACATC 1546
Qy 293 GlnPheLeuMetAlaSerArgAspPheCysValSerPheTyrValIleCysValGluHis 312
Db 1547 GGATTAACAACCTCCAGGTTACCGAGCAGCTAAAGAAATTATGAAAGTCTGTGTGAACAT 1606
Qy 313 HisAlaPhePheArgLeuPheGluGluPheProPheProPheProValLeuPheSer 332
Db 1607 CACACGTTTTCACATTTG--ACATCTACAGACACCATTCCCAAGAAAGCAATTTCTTGCG 1663
Qy 333 ArgGlySerSerPheArgPheSerGlyArgThrGlnValLeuAspTyrValIys 352
Db 1664 CTAGGATCCAAATTCGATACAGTGCAGCTCA----- 1699
Qy 353 GluGlyGlyHisValValGlnPheGluArgHisSerLysIleHisSerIleArg 372
Db 1699 ----- 1699
Qy 373 SerLeuAlaSerGlnProThrGluLeuAsnSerGluValLeuGlnSerGlnGlnSer 392
Db 1700 -----GCTCAGACCGAGGAGCT 1717
Qy 393 ThrSerLeuThrPheGlyGluGlyAlaGluSerProGlyGlnSerCysArgArgGly 412
Db 1718 AGTCCTCTAATT----- 1729
Qy 413 LysGluProLysValSerAlaGlyGluProGlySerHisProSerProAlaProArgArg 432
Db 1730 -----GACAGCGCTGCCCACTTC 1750
Qy 433 SerProAlaGlyAsnValGlnAla-----AspGlyAlaAlaSerAlaProThr 448
Db 1751 GAGCGTACAGCAATTAACGGCGCTCCCGAGCTCGATGAGAGCAGCGCTGC----- 1804
Qy 449 GluGluGluGluGluValValAspArgThrGlnGlnSerLysProGlnProGln 468
Db 1805 -----GATTCGGCAGACCGAAGCTCCG 1828
Qy 469 ProSerThrGlySerLeuThrGlySerProHisLeuSerGluLeuSerValAsnSerGln 488
Db 1829 CCC-----ACTTCGACACCTGCCATTACTCTCAGGCTCAGGTT--GCAGAA 1870
Qy 489 GlyGlyVal-----AlaProAlaAsnValThrLeuSerProAlaLeuSerProAspThr 506
Db 1871 GGTGGCTCTAGATGCTCTGCTAAACAAACAGTGTCTCTAAAGCAGACAGAGAAACA 1930
Qy 507 LysGlnAlaSerProLeuIleSerProLeuLeuAsnArgGlnAlaCysProArgThrAsp 526
Db 1931 GTGAAGCTGAAGTGAAGAAAGAAAGAGCAGCAGCTGAGCAGAGCTGAGCCAGCCACA 1990
Qy 527 AspGluAspGluGlyArgGlyValArgPheProThrAspLysValTyrPheIleAlaLys 546
Db 1991 GAAGCATGAGAAAGAAAGAGAGAAAGACTAGATGATGAGAAACATTTATATC----- 2041
Qy 547 GluValSerThrArgValArgThrTyrLeuLysAsnArgGlnValIleThrSerTyrPhe 566
Db 2042 -----AGACATACGCAATTTAATGTTGAGGATTTAGCAAG-----AGT 2080
Qy 567 GlnSerThrValSerLysGluAspAlaMetProGluAlaLeuLysSerLeuIlePhePro 586
Db 2081 CAAAGAGAGATCAAAAACATCATGCGCAGCATCAGTACGCTGAGAAAG----- 2128
Qy 587 AsnPheGluProLeuHisValArgPheHisThrAsnPheLeuLysGluIle--GluGlnArg 605
Db 2129 -----AATTTCATGAGTCTGACAGAAACAGCG 2158
Qy 606 LeuAlaLeuTyrGluGlyArg 612
Db 2159 CTTAGTGAATGGATTAACGC 2179

```

```

/ Sequence 5, Application US/09664958
/ Patent No. 6916912
/ GENERAL INFORMATION:
/ APPLICANT: Trakht, Ilya
/ APPLICANT: Canfield, Robert
/ APPLICANT: Kalantarov, Gary
/ APPLICANT: Rudchenko, Sergei
/ TITLE OF INVENTION: No. 6916912el Tumor-Associated Marker
/ FILE REFERENCE: 0575/60240
/ CURRENT FILING DATE: 2000-09-18
/ NUMBER OF SEQ ID NOS: 28
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 5
/ LENGTH: 6263
/ TYPE: DNA
/ ORGANISM: Human
/ US-09-664-958-5

Alignment Scores:
Pred. No.: 1,12e-56 Length: 6263
Score: 712.50 Matches: 239
Percent Similarity: 41.5% Conservative: 99
Best Local Similarity: 29.3% Mismatches: 281
Query Match: 13.0% Indels: 198
DB: 3 Gaps: 31

US-09-555-342B-2 (1-1045) x US-09-664-958-5 (1-6263)
Qy 18 GluAsnSerGlyIleSer-----ThrLeuGluArgGlyGlnLysProProProThr 34
Db 369 GAGGCGAGTGGCTTTTCGAGAGAGACCGCCAGAGAGCCAGAAATGCGCCAGAG 428
Qy 35 ProSerGlyLysLeuValSer-----IleLysIleMetLeuAspThrGlnGlu 52
Db 429 ATTGCCAAGAAATACAGAGTGCATTCGCGGCTCATCTGCTTATGCTGCGAGTAT 488
Qy 53 AlaPheGluValProGlnArgAlaProGlyLysValLeuLeuAspAlaValCysAsnHis 72
Db 489 GAGTGTAGGTGAGAGAAACATGCGCGGAGCAGAGTCTGTTTAACTCGTGTGAAC 548
Qy 73 LeuAsnLeuValGluGlyAspTyrPheGlyLeuGluPheProAspHisValIleThr 92
Db 549 CTCACTCTCTAGAGAAAGCATCTTCGCTGACCTTCGTATGCTGACAGCCAGAG 608
Qy 93 ValTyrLeuAspLeuLeuLysProIleValLysGlnIleArg--ArgProLysHisVal 111
Db 609 AACTGGCTGAGACCTCTCCAGAGATCAGAAACAGATCCGAGTAGCCCTCGAATTT 668
Qy 112 ValValLysPheValValLysPhePheProProAspHisThrGlnLeuGlnGluLeu 131
Db 669 GCC-----TTCAAGTCAAGTTCTACCCGCTGATCTCTGCCAGCTGACAGAAACATC 722
Qy 132 ThrArgTyrLeuPheAlaLeuGlnValLysGlnAspLeuAlaGlnGlyArgLeuThrCys 151
Db 723 ACAAGATACACTGTGCTGCGAGCTGCGGAGCATATCATCGGCGCGGCTGCATGC 782
Qy 152 AsnAspThrSerAlaLeuLeuLysSerHisIleValGlnSerGluIleGlyAspPhe 171
Db 783 TCTTTTTCACGAGATGCCCTACATGAGGCTCTCAAGCTGAGCTGAGCTGAGCTGAT 842
Qy 172 AspGluAlaLeuAspArgHis-----LeuAlaLysAsnLysTyrIlePro 187
Db 843 GATGCT-----GAGAGCATGTGGGCAATATGTCAGCAAGCTCCGCTGCCCT 893
Qy 188 GlnGlnAsp---AlaLeuGluAspLysIleValGluPheHisIleAsnHisIleGlyGln 206
Db 894 AACCAAGCCCGGAGAGCTGAGAGAGATCATGAGCTGATTAAGACATATAGAGGATG 953
Qy 207 ThrProAlaGluSerAspPheGlnLeuGlnIleAlaArgArgLeuGluMetTyrGly 226
Db 954 ACCCGGAGAGACAGAAATCACTTCTTAGAGAAATCCAAAGAGCTTTCATATGCGGA 1013

```

Qy 227 lIeargluehIaPfoalalyaAaPArsglUgLyThrylIealnuAlaValaAaen 246
Db 1014 GTAGACTGTCAGCATGCAAGAGCTGAGGGGATGACATCATGTTAGCGTTTGCC 1073
Qy 247 ThrglylIeulValPheglngLyPheThrylIealnuAlaPheantrAlaVal 266
Db 1074 AATGGCTGCTGCTCAACCGGAGCGGCTGAGATCAACGCTTGCTGGCCCAAGATC 1133
Qy 267 ArglyleuSerPheLyAsgLyAsgPheleuIleuLyAsgPheProaAlaAaenSer 286
Db 1134 CTCAGAGTCTCTCAACAGAGAGTAACTTATATCAAGATCCGGCTGGAGGTAG 1193
Qy 287 AlATryAlaPArThryleuglupheleuMetAlaSerAaPheCyAlySerPheTrp 306
Db 1194 CAATTGAGAGCAATGCTTTAGCTTACCTCCAAACACCGGTCAGCAAGACATGTGG 1253
Qy 307 LyelIeCyValaGlnIhIaAlaPhePheAaLygluPheglUgLyProLyPhePro 326
Db 1254 AAGGTGCAATCAAGCATCATCTCTCCGGCTGTG---TCCCTGAGCCCAACC 1310
Qy 327 LyPProValleuPheSerAsgLySerSerPheAaPheSerGlYArThrylInlyglIn 346
Db 1311 AAGGGCTCTGTGTG---ATGGCTCCAGATCCGGTACAGTGGAGAGACCGACAG 1367
Qy 347 ValIeuaPArTyAlaLyglUgLyglYhIeLyelValaGlnPhegluAaLyhIeSer 366
Db 1368 ACTCGCAGGCGCAGCGGCTCATTCAGCCGCTGACCCCTTGTAGCTTGTCCAGC 1427
Qy 367 LyelIeIhIeSerIle---ArgSerLeuAlaSerGlInProThrygluLeuAaenSerGlVal 385
Db 1428 AAGCGTACACATGTCCTCCAGCTTGATGAGACAGATTCCTCCGCCCAAGCTGGTTC 1487
Qy 386 LeuGlIn---SerGlInSerThrySerLeuThryPheglUgLyglYAlaAgluSerProgl 405
Db 1488 AGCGAGACCATGATGAGGGCTGACGGTGAACAGCGGATGAGATGGCGAGTCTGG 1547
Qy 405 YglYglInSerCyAaLyg----- 410
Db 1548 GGGCAACCGTCAGAGGCTGAGAGGAGAGTCAAGCTCAACCAAGATCAAGAGCTA 1607
Qy 411 ---AaGlyLygluPro-----LybVa 417
Db 1608 AAGCCGAGAGGAACACAGCGGACACACAGACAGATTCCTTAAGCAAGCAAGAGAT 1667
Qy 417 lIeA-----AlaGlylupProglYserhIaPfoSerProAlaProAaLySerProAl 435
Db 1668 GTCTTGCTGAGCACAGGCGGATGATGAGCTCAAAAGACCTCGAAGAGCCCAAC 1727
Qy 435 aGlyAaen-----LyglInAlaAaPglYAlaAlaSerAlaProThryglUgI 450
Db 1728 AGCAAACTATCAACCGGATGAGACTGGAAACGGGAGCGGCTGCC----- 1779
Qy 450 uGluGlUgLyValaLyAaPArThryglInSerLyPProglInProglInProse 470
Db 1780 -----CCTCCCGGCTCC-CCCTC 1798
Qy 470 rThryglYserLeuThryglYserProhIeAaenSerGlInLeuSerValAaenSerGlInlygl 490
Db 1799 CCCC-----AAGGGCACCCCTGAG-----AAAGCCATAGAGAGAGG 1837
Qy 490 YAlaAlaProAlaAaenValThryleuSerProAaenLeuSerProAaPArThryglInAlaSe 510
Db 1838 GCTGAGGAGAGGCTCC-----GAGGAGAAAGTCAAAACACACAGTCCCGGGGCC-- 1886
Qy 510 rProleuIleSerProleuLeuAaenPglInAlaCySPArThryThaP----- 526
Db 1887 -----CCAGAGATGACACAGCGCATGA 1909
Qy 527 -AaPgluAaPglUgLyArGArGlyAaPArPheProThAaPlyAlaIaLyPheIleAlaLy 546
Db 1910 GGAACAGAGACAGAGAGGACACAGGCTGTTCTGAGAGACAAACCTGCGCATTTAGCG 1969
Qy 546 aGluValaSerThryThrygluAaPArThryleuLyAaPleuGlU----- 560

Db 1970 CAAGTCTCCAGATCAACGGTCAAGCTCTAGCTGAGGCTGAGAGGCTGAGGTGATTTAC 2029
Qy 561 -ValIeThrySerTrPheglInSerThryValSerlygluAaP-----AlaMetPr 577
Db 2030 GGTCAATGGTGAATCAATGAGGAGCGCTTC-----GAAACTTCTCCGAGCTGCG 2083
Qy 577 aGluAla-----LeuLySerLeuIlePhe----- 585
Db 2084 TGAAGTCGACCGGAGCAAAAGGAGCTCGGACACTAGAGGCTGCTGTTTCCCGGGATCT 2143
Qy 585 ----- 585
Db 2144 CAACAAGGGGCCCCCAGCCAGATGATGATGCTGGGGCATTTAGAGACAGCCCGATG 2203
Qy 586 -----ProAaPheglUgLyPhehIeThryA 597
Db 2204 AAGGGCTGCTCAACCCCGATATGCGCCAGTTTGAAGCCCGTG---AAACAGAAACAT 2260
Qy 597 nPheLeu-----LyglUglInAaLygluAlaLeuTrp----- 609
Db 2261 GACTGCAAGACTTGGCCATTAGAAAGATTAAGACCGGAGCCGCTACTGACAGCAG 2320
Qy 610 ---GlUglYArGSerAaAlaGlnIleAaPArTyglInAaLyglYArValMetIe 628
Db 2321 AGTCTCGCTATGATTAACACACAGAGGTGATGG-AGTCTCTAAGTGGAGGAGT 2379
Qy 628 uLyAaenIleGlnlyMetLyhIeLyAlaAlaIhIeSerTrPlyhIeSerGlAlaIe 648
Db 2380 TCATAGCAACCATCTCCATCAACAGAGACATATGACACACATGAGAAACATTC 2439
Qy 648 uGluAlaLeuGlUaenLy----- 654
Db 2440 TCAAGTCCGGAAAGGGGAGCTGCCATGATCCAGGCCCAAGACGCTGCCAGGAAA 2499
Qy 655 ---IleLySerSerAaLygluAaenPheCyAaPArPheglUglInlybVa 673
Db 2500 TCCGTTCTCTTCTCGATCATCGGAAAGA---TGTCTTCCAGACACTACGGCGCA 2556
Qy 673 lCyAaTyLeuProleuAaenThryPheleuAaLygProleuIhIeLygluMetIeTrly 693
Db 2557 CTGGGAAACCTCTCAACTC----- 2578
Qy 693 aGluValleuGlUaLygluCyAlyhIeIhIaPfoProSerhIa 707
Db 2579 -----CACCAACACCATGTCAAC 2596

RESULT 5
US-09-566-921-57
/ Sequence 57, Application US/09566921
/ Patent No. 6682888
GENERAL INFORMATION:
/ APPLICANT: Loring, Jeanne F.
/ APPLICANT: Tingley, Debora W.
/ APPLICANT: Edwards, Carla M.
/ TITLE OF INVENTION: GENES EXPRESSED IN ALZHEIMER'S DISEASE
/ FILE REFERENCE: PA-0024 US
/ CURRENT APPLICATION NUMBER: US/09/566, 921
/ NUMBER OF SEQ ID NOS: 138
/ SOFTWARE: PERL Program
/ SEQ ID NO 57
/ LENGTH: 6268
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc. feature
/ OTHER INFORMATION: Incyte ID No. 6682888 453592.3
/ NAME/KEY: unsure
/ LOCATION: 2956, 4230-4252, 4825-4887
/ OTHER INFORMATION: a, t, c, g, or other
US-09-566-921-57

Alignment Scores:

Pred. No.:	1,12e-56	Length:	6268
Score:	712.50	Matches:	239
Percent Similarity:	41.54	Conservative:	59
Best Local Similarity:	29.34	Mismatches:	281
Query Match:	13.04	Indels:	198
DB:	3	Gaps:	31

US-09-555-342B-2 (1-1045) x US-09-566-921-57 (1-6268)

```

Qy      18 GluanserGlyIleSer-----ThrLeuGluArgGlyGlnLysProProThr 34
Db      369 GAGGCCGATGGCTTTGCGAGAGACACAGCCAGACAGCCAGAAATCCGCCAGAG 428
Qy      35 ProserGlyLysLeuValSer-----IleYsIleGlnMetLeuAspThrGlnGlu 52
Db      429 ATTCCCAAGAAATACAGAGAGTCCCATCTGCGGGCTGCTGTGAATGCTCGAGAT 488
Qy      53 AlaPheGluValProGlnArgAlaProGlyLysValLeuLeuAspAlaValCysAsnHis 72
Db      489 GAGTGTAGAGTGGAGAAACATGCCGGGGCCAGGTGCTGTTGACCTGTGTGAACAC 548
Qy      73 LeuAsnLeuValGlnGlyAspTyrPheGlyLeuGluPheProAspHisLysLysIleThr 92
Db      549 CTCAACCTCTTAGAAGAGACTACTTGGCTGACCTTGATGTCGACAGCCAGAG 608
Qy      93 ValTrpLeuAspLeuLysProIleValLysGlnIleArg--ArgProLysHisVal 111
Db      609 AACGTGGCTGGACCCCTCCAGAGAGATCAAGAGAGATCCGAGATAGCCCTGGAATTT 668
Qy      112 ValValLysPheValValLysPhePheProProAspHisThrGlnLeuGlnGluLeu 131
Db      669 GCC-----TTCAACAGTCAAGTTTCAACCCGCTGATCTCGCCAGCTGACAGAAAGATC 722
Qy      132 ThrArgTyrLeuPheAlaLeuGlnValLysGlnAspLeuAlaGlnGlyArgLeuThrCys 151
Db      723 ACAAGATACACTAGCTGTGCTGACCTGCGGAGCATCATCAGCGCGGCTGCATGC 782
Qy      152 AsnAspThrSerAlaAlaLeuLeuIleSerHisIleValGlnSerGlnIleGlyAspPhe 171
Db      783 TCTTTGTACAGCATGCGCTTACCTGAGCTCTTACGCTGACGCTGAGCTGGAGACTAT 842
Qy      172 AspGluAlaLeuAspArgGluHis-----LeuAlaLysAsnLysIlePro 187
Db      843 GATGCT-----GAGAGCATGTGGGCAATATGTCAAGCGAGCTCCGCTGCCCT 893
Qy      188 GlnGlnAsp--AlaLeuGluAspLysIleValGlnPheHisIleAsnHisIleGlyGln 206
Db      894 AACCAAGACCGGAGCTGGAGAGAGATCATGAGGTCGATAGACATATAGGGGGATG 953
Qy      207 ThrProAlaGluSerAspPheGlnLeuLeuGlnIleAlaArgArgLeuGluMetTyrGly 226
Db      954 ACCCGGAGAGAGCAAAATCCACTTCTTAGAGATGCCAAGAAAGCTTTCATGTACGGA 1013
Qy      227 IleArgLeuHisProAlaLysAspArgGluGlyThrLysIleAsnLysAlaValAlaAsn 246
Db      1014 GTAACTGTGACCACTGCCAGAGACTGTAGGGCATTCGACATCATGTTAGGGCTTTGGCC 1073
Qy      247 ThrGlyIleLeuValPheGlnGlyPheThrLysIleAsnAlaPheAsnTrpAlaLysVal 266
Db      1074 AATGGCTGCTCATCTACCCGGGACCGGCTGAGATCAACGCTTTGCTGGCCCAAGATC 1133
Qy      267 ArgLysLeuSerPheLysLysLysArgPheLeuIleLysLysLeuArgProAspAlaAsnSer 286
Db      1134 CTCAGATCTCTCTCAAGAGAGATTAATCTTATATCAAGATCCGCGCTGGGAGATATGAG 1193
Qy      287 AlaTyrGlnAspThrLeuGluPheLeuMetAlaSerArgAspPheCysLysSerPheTyr 306
Db      1194 CAATTGTAGACACAAATTTGGCTTTAAGCTCCCAAAACACCGGTGACGCCAAGAGCTGTGG 1253
Qy      307 LysIleCysValGlnHisIleAlaPhePheArgLeuPheGlnGluProLysProLysPro 326
Db      1254 AAGGTCTGATCGAGCATCATATCATTTCTTCCGGCTGGTG--TCCCTGAGGCCCCCAACC 1310

```

```

Qy      327 LysProValLeuPheSerArgGlySerSerPheArgPheSerGlyArgThrGlnLysGln 346
Db      1311 AAGGGCTTCTCGTGG--ATGGGCTCAAGTTCCGGTACAGTGGAGAGACCCAGGCAAG 1367
Qy      347 ValLeuAspTyrValLysGlnGlyGlnHisLysLysValGlnPheGluArgLysIleSer 366
Db      1368 ACTGCCAGGCGACAGGCGCTCATTTGACCGGCTGCACCTTCTTTGAGCGTTTCCAGC 1427
Qy      367 LysIleHisSerIle--ArgSerLeuAlaSerGlnProThrGluLeuAsnSerGluVal 385
Db      1428 AAGCGTACACATGATCCCGCAGCTTGATGACAGACAGATTCTCCGCCAGCTCCGCTC 1487
Qy      386 LeuGlnGln--SerGlnLysSerThrSerLeuThrPheGlyGlnGlyAlaGluSerProGly 405
Db      1488 AGCGAGAACCATGATGTGACGGGCTGACGCTGACAGGAGGAGATGAGATGGCGAGTCTGG 1547
Qy      405 YGlyGlnSerCysArg----- 410
Db      1548 GGGCAAGGTCAAGAGCTGAGAGAGAGAGAGTCCAGATCCAAACGATCAAGAGACTA 1607
Qy      411 ---ArgGlyLysGluPro-----LysVal 417
Db      1608 AAGCGGAGAGAGAAACACAGCCGAGACACAGAGAGATTCTTAGACAAAGCCAGAAAGAT 1667
Qy      417 Lser-----AlaGlyLysProGlySerHisProSerProAlaProArgHisSerProAl 435
Db      1668 GTCTTGCTGAAGCACACAGGCGACGATCATGAGCTCAAAAGACCTCGAAGAGACCCCAAC 1727
Qy      435 ArgLysAsn-----LysGlnAlaAspGlyAlaAlaSerAlaProThrGlnGly 450
Db      1728 AGCAAACTCATCCACGGGAGATGACAGACTGGGAGCGGAGCGCGCT----- 1779
Qy      450 uGluGlnGluValValLysAspArgThrGlnGlnSerLysProGlnProProGlnProse 470
Db      1780 -----CTTCCCGGCTCC-CCCTC 1798
Qy      470 rThnGlySerLeuThrGlySerProHisLeuSerGluLeuSerValAsnSerGlnGly 490
Db      1799 CCCC-----AAGGCAACCCCTGAG-----AAAGCAATGAGAGAGCAGG 1837
Qy      490 YValAlaProAlaAsnValThrLeuSerProAsnLeuSerProAspThrLysGlnAlaSe 510
Db      1838 GCTGAGGAGAGGCTCC-----GAGAGAAAGTCAAAACACACAGTCCCGCGGC-- 1886
Qy      510 rProLeuIleSerProLeuLeuAsnArgGlnAlaCysProArgThrAsp----- 526
Db      1887 -----CCAGAGGTGACACAGCGGATGA 1909
Qy      527 AspGluAspGluGlyArgArgLysArgPheProThrAspLysAlaLysIleAlaLys 546
Db      1910 GAGCCAGAGACAGAGAGAGGACACAGGTGTTCTTGAAAGACAACCACTGCGCATTTGAGCG 1969
Qy      546 GlnValSerThrThrGlnLysArgThrTyrLeuLysAspLeuGlu----- 560
Db      1970 CAAGTGTCCAGCATCAAGCTGACGCTTACGCTGAGCTGAGAGCTGAGGTGACTTCAAC 2029
Qy      561 -ValIleThrSerThrPheGlnSerThrValSerLysGluAsp-----AlaMetPyr 577
Db      2030 GGTCAATTGTGATACATCATGAGCGGCTTC-----GAAGACTTTCCTCCGACGCTGCC 2083
Qy      577 OGluAla-----LeuLysSerLeuIlePhe----- 585
Db      2084 TGAAGCTGACCGGAGCAAAAGCACTGCACTGAGAGGCTGTGTCTTCCGGGATCT 2143
Qy      585 ----- 585
Db      2144 CAACAAGGGGGCCCCAGACGAGATGATGAGTGGGGGATTTGAGGACGCCGGATCG 2203
Qy      586 -----ProAsnPheGluProLeuHisLysPheHisThrAs 597
Db      2204 AAGGGCTGTGTCACACCCCGGATGATGCCCGAGTTTGAGCCCGTG--AAAACAGAAACCAT 2260

```

Qy 597 nphelu-----lysgluilegluigluargleualaleutrp----- 609
Db 2261 GACTGTACAGACTGTGGCCATTAGAAAGAGATTGAGCCGAGCCCTACTGCAGACCAG 2320
Qy 610 ----GluGlyArgSerAsnAlaGlnIleArgAspTyrGlnArgIleGlyAspValMetLe 628
Db 2321 AGTCTCCGATGAGATGACACCCAGCGATTGATGG--GAGTGCCTCAGTGGGAGGAGT 2379
Qy 628 uLyAsnIleGlnGlyMetLeuShiLeuAlaIleHisLeuTrpLysHisSerGluAlaLe 648
Db 2380 TCATGACACCACTCCCTCCATCATCCAGAGACCATTATGACCACTGAGAAACAGTCT 2439
Qy 648 uGluAlaLeuGluAsnGly----- 654
Db 2440 TCAAGTCCGGAGAGGGGCGAGCGTCCATGATCCAGGCCACAGACGCTGGCCACGGAA 2499
Qy 655 ----IleYsSerSerArgArgLeuGluAsnRheCyAsArgAspRheGluLeuGlnLysVa 673
Db 2500 TCCGTTCTCTTTCTCCCATCATCGGGAAGA---TGTCTCACCAGCAGCCTACGGCGCCA 2556
Qy 673 lCySTyrLeuProLeuAsnThrPheLeuLeuArgProLeuHisArgLeuMetHisTyrLys 693
Db 2557 CTGGGAAACCTCTCACTC----- 2578
Qy 693 sGlnValLeuGluArgLeuCyLysHisHisProProSerHis 707
Db 2579 -----CACCAACCACTGTCAC 2596

RESULT 6
US-09-799-451-201

Sequence 201, Application US/09799451

Patent No. 6783969

GENERAL INFORMATION:

APPLICANT: Tang, Y. Tom
APPLICANT: Zhou, ping
APPLICANT: Goodrich, Ryle
APPLICANT: Asundi, Vinod
APPLICANT: Ren, Feiyan
APPLICANT: Zhang, Jie
APPLICANT: Xue, Aidong J.
APPLICANT: Zhao, Qing A.
APPLICANT: Wang, Jian-Rui
APPLICANT: Ma, Yungqing
APPLICANT: Yamazaki, Victoria
APPLICANT: Chen, Rui-hong
APPLICANT: Wang, Ruiwei
APPLICANT: Wang, Dunrui
APPLICANT: Yang, Yonghong
APPLICANT: Wehrman, Tom
APPLICANT: Ghosh, Reena
APPLICANT: Drmanac, Radoje T.
TITLE OF INVENTION: No. 6783969el Nucleic Acids and
FILE REFERENCE: Polypeptides
CURRENT APPLICATION NUMBER: US/09/799,451
CURRENT FILING DATE: 2001-03-05
NUMBER OF SEQ ID NOS: 948
SOFTWARE: pt_FL_genes Version 2.0
SEQ ID NO 201
LENGTH: 3398
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (174)..(2234)
US-09-799-451-201

Alignment Scores:

Pred. No.: 7.2e-49 Length: 3398
Score: 626.00 Matches: 191
Percent Similarity: 44.6% Conservative: 92
Best Local Similarity: 30.1% Mismatches: 221
Query Match: 11.5% Indels: 130

DB: 3 Gaps: 22
US-09-555-342B-2 (1-1045) x US-09-799-451-201 (1-3398)
Qy 31 ProProProThrProSerGlyLysLeuValSerIleLysIleGlnMetLeuAspRhr 50
Db 273 CCGCAGCTGAGGATTCATGATTCATCAGATCAGATGCGGATGCTCTTGTGAGTGTACT 332
Qy 51 GlnGluAlaPheGluValProGlnArgAlaProGlyLysValLeuLeuAspAlaValCys 70
Db 333 GATGTTATGTTGAGCTTCGCAAAAAGCCAAAAGCCAAAGATTTGTTGATTCAGATTATG 392
Qy 71 AsnHisLeuAsnLeuValGluGlyAspTyrPheGlyLeuGlnPheProAsnHisLysLys 90
Db 393 TACCACTGGACCTGATGAAAGCCATATTTTGGTCTGAGATTATGATTCAGACACA 452
Qy 91 lIethValIleTrpLeuAspLeuLysProIleValLysGlnIleArgArgProLysHis 110
Db 453 GTAGCAATTTGGTGTGATGTACAAAAGCATCAAAAAGCAATGAAATTTGGTTCACCC 512
Qy 111 ValValValLysPheValValLysPheProProAsnHisIethGlnLeuGlnGluGlu 130
Db 513 TATTTCTGCATCTTGAAGTAAATTATTTCTCTCAAAACCAATACCTTCGTGAGAG 572
Qy 131 LeuThrArgTyrLeuPheAlaLeuGlnValLysGlnAspLeuAlaGlnIleArgLeuThr 150
Db 573 CTACCCGGATTTATTTGTTCTTCAGTAAACAAAGATATTTCTCAGTGAAATTAAGAC 632
Qy 151 CyAsnAspThrSerAlaAlaLeuLeuIleSerHisIleValGlnSerGluIleGlyAsp 170
Db 633 TGTCCCTTTGATACAGCAGTGCATGCAATTGGCAGCTTAACTCTCAACCTGAATCTGTGAC 692
Qy 171 PheAspGluAlaLeuAspArgLys-----LeuAlaLysAsnLysTyrIle 186
Db 693 TATGATCTTGTCT-----GAGCATGCTCTGAACCTGTCTCAGAGTTCAGATTCGAG 743
Qy 187 Pro--GlnGlnAspAlaLeuGlnAspLysIleValGlnPheHisIleAsnHisIleGly 205
Db 744 CTTATTCAGACTGAGATGAGATGAGATCGCTATTTTGAAGAAATGGAAATGACAGAGGT 803
Qy 206 GlnThrProAlaGlnSerAspRheGlnLeuGluIleAlaArgArgLeuGlnMetTyr 225
Db 804 CAACACCCAGCACAGGCTGAACCAATTCGAATTAAGCCAAAGCGGTAGCAAAAGTAT 863
Qy 226 GlyIleArgLeuHisProAlaLysAspArgGlnGlyThrLysIleAsnAlaValAla 245
Db 864 GGGGTGATATGATGATGTGCTCAAGCTAGAGATGGAAATGATATGATTTGGAGCTAACA 923
Qy 246 AsnThrGlyIleLeuValPheGlnGlyPheThrLysIleAsnAlaPheAsnTrpAlaLys 265
Db 924 CCACAGAGAGTCTTGTGTTTGAAGAGATACCAAAATGGCTTATTTTGGCCGAG 983
Qy 266 ValArgLysLeuSerPheLysArgLysArgPhe--LeuIleLysLeuArgProAla 284
Db 984 ATAAACAGATGGAATTTTGAAGAAATTAATTAACCTGTGTGTGAAGAATGATGAT 1043
Qy 285 AsnSerAlaTyrGlnAsp---ThrLeuGlnPheLeuMetAlaSerArgAspRheCysLys 303
Db 1044 CAGGGCAAGAACAGAGAACATCATTTGCTTTAAGCTGATCATCCAAAGATGCAAA 1103
Qy 304 SerPheTrpLysIleCyAsValGlnHisIleAlaPhePheArgLeuPheGlnGluPro-- 322
Db 1104 CATTTATGAAATGAGCTGTGAGCATCATGCTTCTTCGCGCTT---CGAGGCCCGGTC 1160
Qy 323 ---LysProLysProLysProValLeuPheSerAlaGlySerSerPheArgPheSerGly 341
Db 1161 CAAGAAGTTCATCATGATCAGATGATTTATTCAGATGATCAAGATTAATGATATGAG 1220
Qy 342 ArgThrGlnLysGlnValLeuAspTyrValLysGlnGlyLysIleLysValGlnPhe 361
Db 1221 AAAACGAGATATCAGACCAAAAACCAATAA-----GCAAGAAGATCAACATCTTT 1274
Qy 362 GluArgLysHisSerLysIleHisSerIleArgSerLeu-----AlaSer 376


```

Db      1275 GAAAGAGGCCAGAAAGCATATTCTAGACGAACTCTACAAATGAAGCATGGCTTACA 1334
Qy      377 GlnProthrGluLeuAenSerGluValLeuGlnInserGlnInserTherSerLeuThr 396
Db      1335 AAACCTGMAAGACTT-----AGTGTTCACAAATAAATGTTTCGACC 1373
Qy      397 PheGlyGluGlyAlaGluSerProGlyGlyGlnSerCybaArgArgGlyLysGluProLys 416
Db      1374 CAAAGTAATGGCTCCCAACAGGCTTGGGGAGTAGATGTGCTGCTGCTGAGTCTTCC 1433
Qy      417 ValSerAlaGlyGluProGlySerHisProSerProAlaProArgArgSerProAlaGly 436
Db      1434 ATTTCC----- 1439
Qy      437 AsnLysGlnAlaSerGlyAlaAlaSerAlaProThrGlnGluGlnGluValLys 456
Db      1440 -----TCTGCTCTGTGCGCATGAGATGAG----- 1466
Qy      457 AsparThrGlnGlnInserLysProGlnProGln-----SerProLys 468
Db      1467 -----AATCTTCCACAGAGCTCTGGAACAGACAGCATGACAGAAATGC 1511
Qy      469 ---ProSerThrGlySerLeuThrGlySerProHisLeuSerGluLeuSerValAsnSer 487
Db      1512 ATTCTCTGATATGATGATTGCTGATAGCCACAGACTTATGGAACAGACTT----- 1565
Qy      488 GlnGlyGlyValAlaProAlaAsnValThrLeuSer----- 499
Db      1566 ---GGTGAATTAATGGGGACCTGACACTGAAACATCCAAAGCAGCATGATGACGCTT 1622
Qy      500 -----ProAsnLeuSer-----Pro 504
Db      1623 AATGTAGCCACAGGCTTCCGGAGTAGGGGAACCTGAAGTTGAATAGACACTTAAAA 1682
Qy      505 AspThrLysGlnAla-----SerProLeuLysSerProLeu 516
Db      1683 GACACCTCAGAGAGGCTCAAAACAGCTTGAGATGAGAAACACTCTCTTCTGCTCCCTCGA 1742
Qy      517 LeuAen-----AspGlnAlaCys 522
Db      1743 TCCAACTCATGATGTTAACTAAACAGCAGAGAGAGAGTGTGAAGTTGACTGAGAAATGTC 1802
Qy      523 ProArgThrAspArgGluAspGluGlyArgArg---LysArgPheProThrAsp----- 539
Db      1803 CTTAATTAATGCTCATTAGAGCCACAGATGTAATGTCATGAGATGCTCTGACTTCAAG 1862
Qy      540 -----LysAlaLysPheIleAlaLysGluValSerThrThrGluArgThrTyr 555
Db      1863 AGTAACATTTGAAAGGCTCAAGTAGAAGCACTGCAATAAGTTACAAAGAAAGATAGCTTA 1922
Qy      556 Leu-----LysAspLeuGluValIleThrSerTrpPheGlnSerThrValSerLysGlu 573
Db      1923 TTAAGTCATAAAATGCGCAATGTTCCAGGATGCTGCCAACAACAGTGTGTAAATGAG 1982
Qy      574 AsparLysPro-----GluAlaLeuLysSerLeuIle 584
Db      1983 AATAATGTGCCCCCTCCCAAGAGCTCTGTGAGACTGTGAG 2024

RESULT 7
US-09-906-779-3
; Sequence 3, Application US/09906779
; Patent No. 6770466
; GENERAL INFORMATION:
; APPLICANT: Shi et al.
; TITLE OF INVENTION: Human Protein Tyrosine Phosphatase Polynucleotides, Polypeptides,
; FILE REFERENCE: PT040P1
; CURRENT APPLICATION NUMBER: US/09/906,779
; CURRENT FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: PCT/US01/01563
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/176,306

```

```

; PRIOR FILING DATE: 2000-01-18
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2872
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-906-779-3

Alignment Scores:
Pred. No.: 5,7e-43 Length: 2872
Score: 563.00 Matches: 149
Percent Similarity: 50.8% Conservative: 82
Best Local Similarity: 32.7% Mismatches: 174
Query Match: 10.3% Indels: 51
DB: Gaps: 12

US-09-555-342B-2 (1-1045) x US-09-906-779-3 (1-2872)
Qy      30 LysProProProThrProSerGlyLysLeuValSerIleLysIleGlnMetLeuAspArg 49
Db      6 GAGCGCGCGCGCGCGCGCCAG-GCCACCTCTACTGCGCGCTTCTGCTGAGCGG 64
Qy      50 ThrGlnGluAlaPheGlnValProGlnArgAlaProGlyLysValLeuLeuAspAlaVal 69
Db      65 ACCGAAATGAGCGCGTGAACCTGCCGAAACATGCCAAAGCCAGATTTGTTGATCAAGATT 124
Qy      70 CybaAsnHisLeuAenLeuValGlyLysPheGlyLeuGlnPheProAspHisLys 89
Db      125 GTGTACCACTTGACCTTGTGGAACAGATTAATCTTGCTCCAGTTCTCTGACTGTGCC 184
Qy      90 LysIleThrValTrpLeuAspLeuLeuLysProIleValLysGlnIleArgArgProLys 109
Db      185 CAGGTTCGCGACTGCTGATCATGCGCAACCCATAAAAGAGATGAATAATTGAGCTT 244
Qy      110 HisValAlaValLysPheValValLysPhePheProAspHisThrGlnLeuGlnGlu 129
Db      245 GCTTATGCTTACACTTTCAGATTAATATCATTTCTTCAGAACCAACCACTTCGTGAG 304
Qy      130 GluLeuThrArgTyrLeuPheAlaLeuGlnValLysGlnAspLeuAlaGlnGlyArgLeu 149
Db      305 GAGTTTCAAGAGTACCTGTTGTTTCAACTCAGCATGACATCTTCTGGAATAATTG 364
Qy      150 ThrCybaAsnAspThrSerAlaAlaLeuLeuIleSerHisIleValGlnInserGlnIleGly 169
Db      365 AAATGCCCTTAATGAACAGCTGTGAAATAGCTGCTCTGTCTACAAAGGAGGACTGGG 424
Qy      170 AspPheAspGluAlaLeuAspArgLysHis-----LeuAlaLysAsnLysTyr 185
Db      425 GAGTGCAGACTTCCA-----GAACACACACACAGCATGCTGTGAGTTCCGCTTC 475
Qy      486 IleProGlnGln---AspAlaLeuGluAspLysIleValGluPheHisAsnHisIle 204
Db      476 ATTCCAATGACACAGAAAGCAATGGAATTGATTTCTTCACAGATGGAAGAGTGCAGG 535
Qy      205 GlyGlnThrProAlaGluSerAspPheGlnLeuLeuIleAlaArgGluLeuMet 224
Db      536 GGAAGAGCCCTGCGCAGGCGAACTCTCTATGTGAATAAAGGAAAGTGCCTGGAATG 595
Qy      225 TyrGlyIleArgLysHisProAlaLysAspArgGluGlyThrLysIleAsnLeuAlaVal 244
Db      596 TATGGGTTAGACATGCACCTGTCAGGGGAAGATGCGTGTGAATTTCTCTTGGACTG 655
Qy      245 AlaAsnThrGlyLysLeuValPheGlnGlyPheThrLysIleAsnAlaPheAsnTrpAla 264
Db      656 ACCCGCAGAGCATTAATTAATCTTGAAGAGCTPAACAAATAGGCTTAATTTTGCCCT 715
Qy      265 LysValArgLysLeuSerPheLysArgLysArgPhe---LeuLysLysLeuArgProAsp 283
Db      716 AAATTTACAAATAGATTTTAAAGAGCAAAATGACACTCGGTGCTGAGAGATAT 775
Qy      284 AlaAsnSerAlaTyrGlnAsp---ThrLeuGluPheLeuMetAlaSerArgAspPheCys 302

```

Db 776 GATCAGGAGCGTGAAGACACACCTTTGTGTTCCGTTAGACAGTGCAGACCTGC 835
Qy 303 LysSerPheTrpLysIleCysValGluHisAlaPhePheArgLeuPheGluPro 322
Db 836 AAACACCTTTGGAAAGTGTGCAAGTTGAGCACACGCAATTTCCGACTGCCGAGCCAGCA 895
Qy 323 LysProLysProLysProValLeuPheSerArg---GlySerSerPheArgPheSerGly 341
Db 896 AACAGCAATTCATATAGATCCGACTTATCAGAGGTGGCTCTGGCTTCAGATTCAGTGG 955
Qy 342 ArgThrGlnLysGlnValLeuAspTyrValLysGluGly-----HisLysLysVal 359
Db 956 CGGACGAAATATCA-----GCTACACATGCTCCAGGTTACGAAAGACCGC 1003
Qy 360 GlnPheGluArgLysHisSerLysIleHisSerIleArgSer-----LeuAla 375
Db 1004 ACCTTTGAGAGAAAGCTTAGTAACGTTATCCATCCCGAGACATTCACGTTCAAGCA 1063
Qy 376 SerGlnPro-----ThrGluLeuAsnSerGluValLeuGluGln 388
Db 1064 AGCAACCCAGTATAGACAGCCAGCTGCTCTAAACAAATCCAGAAATCCATAATTAC 1123
Qy 389 SerGlnLysSerThrSerLeuThrPheGlyGluGlyAlaGluSerProGlyGlyGlnSer 408
Db 1124 CAGCTTCA----- 1132
Qy 409 CysArgArgGlyLysGluProLysValSerAlaGlyGluProGlySerHisPro---Ser 427
Db 1133 -----TATCATCTTAATATCCATCCAGCCAGCCCGGTGGCATCTCTACTCT 1180
Qy 428 ProAlaProArgArgSerProAlaGlyAsnLysGlnAlaAspGlyAlaAlaSerAlaPro 447
Db 1181 CCAATGTCAGGCGCATCTTTCAGAGATGACAGGTCGATGGAAGATCGGCGACATGGA 1240
Qy 448 ThrGluGluGluGluValValLysAspArgThrGlnGlnSer 462
Db 1241 GATGACAGCCATTTGATGTATGTCCAGACGAAACGAAAGAAC 1285
RESULT 8
US-10-104-047-41
Sequence 41, Application US/10104047
Patent No. 6943241
GENERAL INFORMATION:
APPLICANT: HELIX RESEARCH INSTITUTE
TITLE OF INVENTION: No. 6943241el full length cDNA
FILE REFERENCE: H1-A0105
CURRENT APPLICATION NUMBER: US/10/104,047
PRIOR FILING DATE: 2002-03-25
NUMBER OF SEQ ID NOS: 4096
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 41
LENGTH: 2493
TYPE: DNA
ORGANISM: Homo sapiens
US-10-104-047-41
Alignment Scores:
Pred. No.: 1,49e-42 Length: 2493
Score: 557.50 Matches: 200
Percent Similarity: 43.5% Conserved: 103
Best Local Similarity: 28.7% Mismatches: 248
Query Match: 10.2% Indels: 146
DB: 3 Gaps: 27
US-09-555-342B-2 (1-1045) x US-10-104-047-41 (1-2493)
Qy 8 ProThrProGly-----SerArgLeuGlyAlaProGlyAsnSer 20
Db 163 CCTGGTCACAGGTCGCCGAGCGCCGCGCTCGCTCCCGGCGGCGGCGGAGATGCT 222
Qy 21 GlyIleSerThrLeuArgLysGlyGlnLysProProThrProSerGlyLysLeuVal 40

Db 223 GAG---CAGTTGATGACGCGGAGCAGCAGAGCT-----GAGCGCGAGTA 267
Qy 41 -SerIleLysIleGlnMetLeuAspArgThrGlnGlnAlaPheGluValProGlnArgAl 60
Db 268 CAGCTCAGCCGTCGCGGCTCTGAGCAGCAGAGTACACCTTGACCATCCAGCAAGATGC 327
Qy 60 aProGlyLysValLeuLeuAspAlaValCysAsnHisLeuAsnLeuValGluLysArg 80
Db 328 CAAAGCCAGTACTCTTTGACCTCTTGGCCACCATCTGAACTCTTGAGAAACATA 387
Qy 80 rPheGlyLeuGluPheProAspHisLysLysIleThrValTrpLeuAspLeuLysPr 100
Db 388 TTTGGTATCCGCTTGTAGACCCAGATAGACGCGCATGGCTGCAATTT----- 439
Qy 100 oIleValLysGlnIleArgArgProLysHisValValLysPheValValLysPhe 120
Db 440 -----ACAAAGTCGTGTGTAAGTTTA 462
Qy 120 eProProAspHisThrGlnLeuGlnGluGluLeuThrArgTyrLeuPheAlaLeuGlnVa 140
Db 463 TCCTGACAGCCCTGCTGCTGTGAAGAAATAACACAGTATTTAGTCTCTCGAGAT 522
Qy 140 LysGlnAspLeuAlaGlnGlyArgLeuThrCysAsnAspThrSerAlaAlaLeuLeu 160
Db 523 CAAAGGATCTCTACATGCGCCGACTCTCTGTAAACATGAGATGCTGCTTTAGC 582
Qy 160 eSerHisIleValGlnSerGluIleGlyAspPheArgLysAlaLeuAspArgLys---His 179
Db 583 AGCTTACATCTTCAAGCGAAGATGGGATTAATAGCTCAGAGAAACACCTGAAGCTTA 642
Qy 179 sLeuAlaLysAsnLysTyrIleProGln---GlnAspAlaLeuGluAspLysIleValG 198
Db 643 CAGCTCCAGTTCACGTTTCCCTTAACATTCAGAGAGCTGGAAGAAATTCCTGA 702
Qy 198 rPheHisHisAsnHisIle---GlyGlnThrProAlaGluSerAspPheGlnLeuGlu 217
Db 703 GATTCACAGACGGAAGTAGTGTCAACACACAGACATCATGAGCTGAATCTTCTTAAG 762
Qy 217 uIleAlaArgArgLeuGluMetTyrGlyIleArgLeuHisProAlaLysAspArgLys 237
Db 763 AAAAGCAGACATTTGGAACATATGAGTGAATCTCCACCAATGTATAGAGAGCTGACAG 822
Qy 237 yThrLysIleAsnLeuAlaValAlaAsnThrGlyIleLeuValPheGlnGlyPheThrLys 257
Db 823 AATGTCGATTTCTGCGCTTCACTCCCTTTGGTTGTGTTCTTCAAGAAACAAGAG 882
Qy 257 sIleAsnAlaPheAsnThrAlaLysValArgLysLeuSerPheLysValGlyArgPheLe 277
Db 883 GGTCCACTTCATTAATGAAATGAGGTGACCAAGCTGAATTTGAAGGAAGACTTTCTTA 942
Qy 277 uIleLysLeuArgProAspAlaAsnSerAlaTyrGlnAspThrLeuGluPheLeuMetAl 297
Db 943 TTATATCGTA-----AGTCAGAAAGAGAAAGAAATTAATTTCTTACATATTTTCTCC 996
Qy 297 aSerArgAspPheCysLysSerPheTrpLysIleCysValGluHisAlaPhePheArg 317
Db 997 AACTCCTGAAGGTGTGAAGCAGCTCTGAAATGTGAATCGAAGAACCAAGCTTCTACAA 1056
Qy 317 GluPheGluGluProLysProLysProLysProVal-----LeuPheSerArg 333
Db 1057 GCTG-----GAGAGTCAGCCCAAGTCCGACAGTGTCCAGCAGCAATTTATTTCTTAA 1110
Qy 333 gGlySerSerPheArgPheSerGlyArgThrGlnLysGlnValLeuAspTyrValLysG 353
Db 1111 AGGAGACCGCTTCCATACAGTGGCCGAGCTTCAAGGAAGATCATGGA----- 1159
Qy 353 uGlyGlyHisLysLysValGlnPheGluArgLysHisSerLysIleHis----- 369
Db 1160 -----TCAAGTCTAATATCAACGGAGACACCGGAATATCAAGAGCAGGAGT 1209
Qy 370 -----SerIleArgSerLeuAlaSerGlnProThrGluLeuAsnSerGluVa 385

```
D 1210 GATTCCAGCCGAGCTGTCCCTCCATTAACCATGAGCCCAAG-----CTGAGCAGGT 1263
Q 385 |LleuGlInSerGlnInSerThrSerLeuThrPheGlyGluGlyAlaGlySer----- 403
D 1264 CCCAGAGACCCGAGAAAGCTGTTCATCTCATATGAGAGGCTTAGAGCTTAAAG 1333
Q 404 -----ProG1 405
D 1324 GAGAGTCCCATTCACACCAAGTCCGTTCACATTCCTCCATGGGAGACCTTCGCTCA 1383
Q 405 yGlyGlnSerCyArGArgGlyGlyGluProLyValSer-----AlaG1 420
D 1384 CGTGAGAGGACCGCGACAGATAGCAATGAGAGTAGAGTGTGATGCGAGCAGGCTTA 1443
Q 420 yGluProGlySerHis-----ProSerProAlaProArgArgSer-----ProAl 435
D 1444 CAGCCCTGCAGACAGCCTGCTGCTCCACCTGTGGCTGAGCAGACGCTGAGTGT 1503
Q 435 aGlyAsnLyGlnAlaAspGlyAla-----Al 444
D 1504 GCTTCCCGGCAATCATAGGAGCCACCTGCAGCATTTGAGAGAGAGAAAGAAATCTGAAC 1563
Q 444 aSerAlaProThrGlnGluGluGluValValLyAspArgThrGln----- 460
D 1564 CAGACCCCAACTGCTACAGAGGTGAGAGCCCTGGGGAGAGCTGAGAGGCTGTGTCA 1623
Q 461 -----GlnSerLyAspProGlnProProGlnProSerThrG1SerLeuThr-GlySer--- 477
D 1624 GGGGACAGCGGGCCGAGAGAGAAAGAGTGAATGATTGTTTGAAGTGTCTCCGTTT 1683
Q 478 -----ProHisLeuSerGluLeuSerValAsnSerGlnGlyGlyValAlaProA 494
D 1684 GCTCTCTGTGACCAATGGAGACTCTCTTGTGTTGCTCTCTCTCTGATCATCTTACCG- 1742
Q 494 |aAsnValThrLeuSer-----ProAsnLeuSerP 504
D 1743 --AGTGCAGCTTGCATATGCTTTTCCGTGATATCCGACAGACCCCGAGTTTGAACA 1800
Q 504 roAspThr-----LyGlnAlaSerP 511
D 1801 ATTCACATATCAATATCTTTGTCCCTCAGCGCATGCTTGGCTGCAGAAATCCCTCAGT 1860
Q 511 roLeuHisSerProLeuLeuAsnAspGlnAla-----CyProArgThrAspAspGluA 529
D 1861 GGTGAGCCTGCTATTCACACCTGAGAAAGCATGATCTTCCAAAACTHAGCCAGGTG 1920
Q 529 ePrlGluGlyArgGlyAspArgPheProThrAspLyValATyrPheHisAlaLyGluValS 549
D 1921 ACCAAGAACCCGCGCTACCCATTCGCCAGCAATGGAGCCCAT-----C 1962
Q 549 eArThrThrGluArgThrTyrLeuLyAspLeuGluValHisThrSerTrpPheGlnSerT 569
D 1963 GCGGAAACCATCGGACATATACCAAGTCTCTCTCATGACTCA-----AAGTCCA 2013
Q 569 hrValSerLyGluAspAlaSerProGluAlaLeuLySerLeuHisPheProAsnPhg 589
D 2014 CTGCAAGCTTGAAGAGG-TGTTTCCCAAGAAAGAAAGAGATGAGCTATGCTGTCTAA 2072
Q 589 |uProLeuHisLyS-----PheHisThrAsnPheLeuLyGlu 601
D 2073 ACAAACTGGGAAATCATATTTCTTCAAGAAATTATTTCAAGAA 2116

RESULT 9
US-09-848-294-1
/ Sequence 1, Application US/09848294
/ Patent No. 6479640
/ GENERAL INFORMATION:
/ APPLICANT: Tokes, Nicholas K.
/ TITLE OF INVENTION: Isolation of A cDNA Encoding A No. 6479640el
/ TITLE OF INVENTION: Protein Tyrosine Phosphotase Which Localizes to Focal
/ FILE REFERENCE: CSHL90-04FZA
/ CURRENT APPLICATION NUMBER: US/09/848,294
```

```

/ CURRENT FILING DATE: 2001-05-03
/ PRIOR APPLICATION NUMBER: 09/235,251
/ PRIOR FILING DATE: 1999-01-22
/ PRIOR APPLICATION NUMBER: 08/759,536
/ PRIOR FILING DATE: 1996-12-04
/ PRIOR APPLICATION NUMBER: 08/107,420
/ PRIOR FILING DATE: 1993-08-16
/ PRIOR APPLICATION NUMBER: 07/663,579
/ PRIOR FILING DATE: 1991-03-01
/ PRIOR APPLICATION NUMBER: 07/494,036
/ PRIOR FILING DATE: 1990-03-14
/ SOFTWARE: FASTSEQ for Windows Version 3.0
/ NUMBER OF SEQ ID NOS: 13
/ SEQ ID NO 1
/ LENGTH: 3984
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (24)...(2765)
US-09-848-294-1

Alignment Scores:
Pred. No.: 1,29e-40 Length: 3984
Score: 541.00 Matches: 229
Percent Similarity: 40.1% Conservative: 151
Best Local Similarity: 24.2% Mismatches: 358
Query Match: 9.9% Indels: 210
DB: Gaps: 35

US-09-555-342B-2 (1-1045) x US-09-848-294-1 (1-3984)
Q 27 ArgGlyGlnLyAspProProProThrProSerGlyLyLeuValSerLeuHisGlnMet 46
D 69 CGCACTCGGAGTAACTCCCAAGAAATCGATCAGAAAGCTATTTGGCTTACAGCATTT 128
Q 47 LeuAspAspThrGlnGluAlaPheGluValProGlnArgAlaProGlyLyValLeu 66
D 129 TTGATGGCGGTGACAGACCTTTAAAGTTACTTAAACAAACATGCGCCAGTTCTTG 188
Q 67 AspAlaValCyAsnHisLeuAsnLeuValGluGlyAspTyrPheGlyLeuGluPhePro 86
D 189 GATATGTGCACAACACCACTGGGTGTGACTGAAAGAGATATTTGGTTTACAGCATAT 248
Q 87 AspHisLyValSerLeuThrVal---TrpLeuAspLeuLeuLySerProHisGlnHis 105
D 249 GAGGACTCCGTGACCTCTCTAGATGCTGGAAGCAAGCAAACTCAGAAAGCAGTTA 308
Q 106 ArgArgProLyHisValValValLyPheValValLyAspPheProProAspHisThr 125
D 309 AAGGAGGTTTCCCTGTACCTGATTTTCAGATTAAGATTTTATTAATCTGATCCCAAC 368
Q 126 GlnLeuGlnGluGluLeuThrArgTyrLeuPheAlaLeuGlnValLyGlnAspLeuAla 145
D 369 ACACTGAGCAAGAACAAACACGAGCACTTGTATTTCTTACACTGAAAGATGATATTGC 428
Q 146 GlnGlyArgLeuThrCyAsnAspThrSerAlaAlaLeuLeuHisSerHisGln 165
D 429 GAAGGAAGTTAACTGCTCTTAACTCAGAGTGTCTACAGTCTTACGCTTACGCTTAA 488
Q 166 SerGlnLyAspPheAspGluAlaLeuAspArgGlu---HisLeuAlaLyAsnLyS 184
D 489 TTTCAATTTGAACTATATATTTCTTCATACATATCCAGGCTATTTTCGATAGTCAC 548
Q 185 TyrHisProGlnHisAspAlaLeuGluAsp-----LyHisLeuAluPheHisHis 201
D 549 TTTATACCCGATCAAAAT-----GAGGACTTTTAAACAAAGTCGAAATCTGATGAG 602
Q 202 AsnHisGlnGlyGlnThrProAlaGluSerAspPheGlnLeuLeuGlnHisAlaArg 221
D 603 CAGCAGTGTGTGTTAAACATACAGAAAGCAATCTGCTATTTCAACATAGCGCGAGCC 662
Q 222 LeuGlnMetTyrGlyLeuArgLeuHisProAlaLyAspArgGluGlyThrLySLeu 241
```



```

Qy      910 rmevalhivalcysrttr 916
Db      2433 -CTCAGTACGTCCGATGG 2450

RESULT 10
US-09-949-016-3593
; Sequence 3593, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CU001307
; CURRENT APPLICATION NUMBER: US/09/949, 016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 3593
; LENGTH: 2156
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-3593

Alignment Scores:
Pred. No.:      6.28e-38      Length:      2156
Score:          508.00      Matches:      148
Percent Similarity: 47.0%      Conservative: 78
Best Local Similarity: 30.8%      Mismatches: 198
Query Match:      9.3%      Indels:      57
                        Gaps:      14

US-09-555-342B-2 (1-1045) x US-09-949-016-3593 (1-2156)
Qy      27 ArgGlyIuInyPrProProThrProSerGlyLysLeuValSerIleLysIleGlnMet 46
Db      63 CGCACCTCGGAGTACCCAAAGAGAAACTCGATCAGAGTCATTGGACATCCACTT 122
Qy      47 LeuAspThrGlnGlnuAlaPheGluValProGlnArgAlaProGlyLysValLeuLeu 66
Db      123 TTGATGGCGTGTGACAGACCTTAAAGTTACTAAACAGACCTGGCCAGGTTCTTCTG 182
Qy      67 AspAlaValCysAsnHisLeuAsnLeuValGluGlyAspTrpPheGlyLeuGlnPhePro 86
Db      183 GATATGGTGCAACAACCTGGGTGTGACTGAAAAGAAATATTTGGTTTACAGCATGAT 242
Qy      87 AspHisLeuValSerIleThrVal---TrpLeuAspLeuLeuLysProIleValLysGlnIle 105
Db      243 GAGCAGCTCCGTCGACTCTCTCAGATGGCTGGAGCAAGCAAGCAATCAGAGACAGTTA 302
Qy      106 ArgArgProLysHisValValValLysPheValValLysPhePheProProAspHisThr 125
Db      303 AAAGGAGTTTCCCTGATCCCTGATCTTTCGATTAAGATTTTTAATACCTGATCCCAAC 362
Qy      126 GlnLeuGlnGluGluLeuThrArgTrpLeuPheAlaLeuGlnValLysGlnAspLeuAla 145
Db      363 ACATGCGAGCAAGAACAAACCAAGGCACTGTATTTCTTACAACTGAAGATGATATTTGC 422
Qy      146 GlnGlyArgLeuThrCysAsnAspThrSerAlaAlaLeuLeuIleSerHisIleValGln 165
Db      423 GAAGGAGGTTTAACTGCTCTTAACCTCAGACAGGTTTAAAGGTCCTTAAGCCGTACAA 482
Qy      166 SerGluIleGlyAspPheAspGluAlaLeuAspArgGlu---HisLeuAlaLysAsnLys 184
Db      483 TCTCATTTTGGAGACTATTAATTTCTTCATATCATTCAGGAGCTATTTTTCGATAGTCAC 542
Qy      185 TyrIleProGlnGlnAspAlaLeuGlnAsp-----LysIleValGluPheHisHis 201

```

```

Db      543 TTTATACCCATCAAAAT-----GAGGACTTTTAAACAAAGTCATCTGATGAG 596
Qy      202 AsnHisIleGlyGlnThrProAlaGluSerAspPheGlnLeuGlnIleAlaArg 221
Db      597 CAGCAGTGGGCTTAAACAAACAGAGCAAGAACTCTGTATATCAATATGACCGGAGCC 656
Qy      222 LeuGlnMetCysGlyIleArgLeuHisProAlaLysAspArgGluGlyThrLysIleAsn 241
Db      657 CTCGACTTCTATGAGAAAGTGAACAGTGGTAGGATCTGCACAAATTTAGACTATAG 716
Qy      242 LeuAlaValAlaAsnThrGlyIleLeuValPheGlnGlyPheThrLysIleAsnAlaPhe 261
Db      717 ATTGAATTCCTTCCGGGGTCTGCTGTGTGACCAAGAAATACATTTCACAAAGTTCTAT 776
Qy      262 AsnTrpAlaLysValArgLysLeuSerPheLysArgLysArgPheLysIleLysLeuArg 281
Db      777 CTTGGGTGAACATCTCAAAATTTCTTCAAAAGGAAAAAGTTCTTCATCATCAGCGA 836
Qy      282 ProAspAlaAsnSerAlaTrpGlnAspThrLeuGlnPheLeuMetAlaSerArgAspPhe 301
Db      837 CAGAAACAGCTGAATCAGGAAACATATTTGGCTTCAACATGCTGAATTACCGATCT 896
Qy      302 CysLysSerPheTrpLysIleCysValGlnHisIleAlaPhePheArgLeuPheGluGlu 321
Db      897 TGCAAAACCTTGTGAAATCTGTGTGACACCATACGTTCTTTCAGGCA---AAGAG 953
Qy      322 ProLysProLysProLysProValLeu-----PheSerArgLysSerPheArg 338
Db      954 CTACTACCTCAGGAAAGATGTTCTGTCTCAGTACGACTAGGCTCTCGG----- 1007
Qy      339 PheSerLysArgThrGlnLysGlnValLeuAsp---TyrValLysGlu-----GlyGly 355
Db      1008 -----AACCCAAAAGTCGTAATATACCAATATTCGAAAAAGTGATGGGGG 1058
Qy      356 -----HisLysValGlnPheGluArgLysHisSerLysIle 368
Db      1059 ATGCTGGAACCCAGCCATCGGAGATCCTTATCAGTGGAGCACTTAGAAACCAAGT 1118
Qy      369 HisSerIleArgSer-LeuAlaSerGlnProThrGlnLeuAsnSerGluValLeuGluG 388
Db      1119 CTGCTCTCTGTTCCCTCCCATTAATCTCCCACT----- 1152
Qy      388 nSerGlnGlnSerThrSerLeuThrPheGlyGluGlyAlaGluSerProGlyGlnIle 408
Db      1153 -----GGGAGAGTGGATCAAAAGATGCCTC 1178
Qy      408 rCysArgArgGlyLysGluProLysValSerAlaGlyLysProGlySerHisProSerPr 428
Db      1179 TTGTGATATCAAGATTAACCCAGATCAGCTCGGACAC-----CTGCAT---TCTTA 1229
Qy      428 AlaProArgArgSerProAlaGlyAsnLysGlnAlaAspGlyAlaAlaSerAlaProTh 448
Db      1230 AGCTGAACGAAGGGGATCAATCGTTTAATCAATGCGCGGACATCTCAGAACACAGC 1289
Qy      448 rGluGlnGluGlnValValLysAspArgThrGlnGlnSerLysProGlnProProG 468
Db      1290 ATGACCAAGTGTGATGTT-----CATTAAGCCAGCCGGAGT 1328
Qy      468 nProSerThrGlySerLeuThrGlySerProHisLeuSerGluLeuSerValAsnSerG 488
Db      1329 CCCACTACCGGAGCTGCGCTGTATCAGAGGAGACGACTCTACAGAAAAAAGCCA 1388
Qy      488 n 488
Db      1389 G 1389

RESULT 11
US-09-270-767-1995
; Sequence 1995, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.

```


FILE REFERENCE: 084335/0160
CURRENT APPLICATION NUMBER: US/10/094,749
CURRENT FILING DATE: 2002-03-12
PRIORITY APPLICATION NUMBER: 60/350,435
PRIORITY FILING DATE: 2002-01-24
PRIORITY APPLICATION NUMBER: JP 2001-328381
PRIORITY FILING DATE: 2001-09-14
NUMBER OF SEQ ID NOS: 3381
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1401
LENGTH: 2931
TYPE: DNA
ORGANISM: Homo sapiens
US-10-094-749-1401

Alignment Scores:
Pred. No.: 1,09e-36 Length: 2931
Score: 497.50 Matches: 207
Percent Similarity: 41.3% Conservative: 161
Best Local Similarity: 23.2% Mismatches: 377
Query Match: 9.1% Indels: 146
DB: 4 Gaps: 31

US-09-555-342b-2 (1-1045) x US-10-094-749-1401 (1-2931)

QY 248 G1Y1LeuValPheGlnGlyPheThrIleuValPheAen----- 262
DB 85 GGCCTGGCTGGCTCGCTGGAAGAAACAATGAGAGCTTTTACTTCAACCATTTGTGAC 144
QY 263 ---TrrAlaValAlaGlyLeuSerPheValGlyValGlyPheLeuIleValLeuArg 281
DB 145 TTCTGGCCAGAGAAATAATTCAGCTTTTAAAGCAGCAGTACTGTCCAAAGATGCG 204
QY 282 ProAerAlaAenSerAlaTyrglnAspThrIleuGlnPheLeuMetAlaSer----- 298
DB 205 -----TTTAGTTCACCTTGGCTCCACAGACAGACACACACATGCTGT 246
QY 299 -----ArgAspPheCysLeuSerPheTrpIleCys-ValGlnI 312
DB 247 TGGTGAATGTATCTGAAGAAAGGCTCAGAGGATAAATGGGAACAGCCAGCAACA 306
QY 312 vHValPhePheArgLeuPheGlnGlyProIleValGlyValGlnPheGlnI 330
DB 307 CTGACGT-----GCAAGTCCAAACCCACAACTGCTCCAAAGCCATTACA 351
QY 330 uPhe-----SerArgIleSerSerPheArgPheSerGlyArg-----Th 343
DB 352 CCTGCAAAATTCACCTTCCTCATATATACACAAACCCCGAGCATAAAGCTTAACCTAG 411
QY 343 rGlnIleuGlnValLeuAspTrpValIleGlnGlyIleValIleValIleValIle 363
DB 412 TGCAAAACCAAGATGAGAGAAATTAACCTGCTGCTTGTGTCTCAAAAGAA-- 469
QY 363 gLValHISerIleValIleHISerIleArgSerIleValIleSerGlnProThrGlnLeuAen 383
DB 470 -AAACCAAGTAAATGATCATCATCATGCTGCTTGAAGAGAGCAGCTATTATCAA 528
QY 383 rGlnValIleuGlnIleuSerGlnIleuSerIleuThrPheGlnIleuValIle 403
DB 529 TTATAGATTTTGAAGAAAGCTGCTGTAACCTA-----AATGCTCTTAAGAC 579
QY 403 rProGlyGlyIleuSerCysArgArgGlyIleGlnProIleValSerAlaGlyIle 422
DB 580 CCCAGAAAGGATGATGACACACACCTCAACAAATCTCTCCAGCACTTGCC 639
QY 422 oGlySerHISerProAlaProAlaProArgArgSerProAlaGlyAenIleValIle 442
DB 640 ACAAGGCAAGGAAATGATGCAATTAAGATCAGAGGTGCAAGACTGTGTGGCAACG 699
QY 442 yAlaAlaSerAlaProThrGlnIleuGlnIleuValIleValIleValIleValIle 462
DB 700 TGTAAATGAGCAACAAACCAAGATGATGAGAGAGAAAGCTCCACTTAACTC 759

QY 462 rLysProGlnProProGlnProSerThrGlySerLeuThrGlySerProHISerLeuSerG 482
DB 760 A-----GATACTCTTATTAAGCTTCAACCTTGTGATAC 798
QY 482 uLeuSerValAsn-----SerGlnGlyValAlaProAlaAenValThrLeuSe 499
DB 799 GCACATATGTAATGAGAAAGATGAATGCAAGCTCTGTCATCACCCACAAACAGA 858
QY 499 rPro-----AsnLeuSerProAspThrIleGlnAlaSerProLeuIleSerProLe 516
DB 859 TAGCTGTATGAAATCTCT--GACAGTACGTACAGATCTCCAGCATAGGCCAGT 915
QY 516 uLeu-----AsnAspGlnAlaCysProArgThrAspAspGlnAspGlnI 531
DB 916 GCTCCCTTGAAGAAAGAGGGGCGAAACAGAAACCAAGGTACAGAGGAAATGG 975
QY 531 Y-----ArgArgIleValArgPheProth 538
DB 976 GAAAGCCCTGTGAATGTGAGCAGCTGACACAGACCATGATGAAGAGATTAATGA 1035
QY 538 rAspIleValIleThrPheIleAlaIleGlyValIleSerThrGlnArgThrTy 558
DB 1036 GCAAAATTCACAAATATGACATGAACTTTGCTTACTGAAGACCTTATGCAACG 1095
QY 558 rLeuGlnValIleThrSerTrpPheGlnIleuSerThrValSerIleGlu-----AspAl 575
DB 1096 ACTGACCTCTTGATAGATGATTTTATGCAACCTTGGAAACAAACCAAGCTC 1155
QY 575 aMetProGlnAlaLeuIleuSerLeuIlePheProAsnPheGlnProLeuHISerPheH 595
DB 1156 GTTTCAGCAGAGATGTGTAATAATCTTTCCAAATTTTCATCAATAATGCTTCCA 1215
QY 595 vThrAspPhe-----LeuIleGlnIleGlnIleValIleuThrGlnIleValIle 614
DB 1216 TAGTAAATCTCTTGCACAGCTGGAACCAATGCAATGAGGAA----- 1264
QY 614 nAlaGlnIleArgAspTrpGlnIleGlnIleValIleMetLeuIleValIleGlnIle 634
DB 1265 -----ACTACTCTTGAATTTGAAGATCTTCAAGAAATTTGGACCAATTCCT 1311
QY 634 tLysHISerLeuAlaAlaHISerTrpIleHISerGlnAlaLeuGlnIleuGlnI 654
DB 1312 TAAAGTATGAGAAATATGTAAGAGATTTGATTAATGCAATGAAATGTTGTAATAACAT 1371
QY 654 yIleIleSerSerArgArgIleGlnIleuAsnPheCysValGAspPheGlnIleuGlnI 674
DB 1372 GACGAAACGTATTTCCCAAGTTCAAATGAGGTGTAAGAAATTCAGAAACAGAAATCTG 1431
QY 674 s-----TyrlleuProLeuAsnThrPheLeuIleuArgProLeuHISerGlnIle 693
DB 1432 TGGAGCTTAACCTTGAAGATCAATGATGAAACCTGTTCAGGAGATTCCTCCGATAGA 1491
QY 693 gGlnValIleuGlnIleuGlnIleuValIleProProSerHISerAlaAspPheArgSer 713
DB 1492 GATGCTCTTAAGACATATGTAAGAAATGCTCTGATTCCTGAGTCAAGATGATG 1551
QY 713 vArgAlaAlaLeuAlaGlnIleThrGlnIleuSerValAlaGlnIleuHISerIle 733
DB 1552 TAAAAATTCACCTTGAATTTATCTACAGACAGCAAGCAATTTGATGTCATTAAGAA 1611
QY 733 vMetGlnAspPheGlnIleuValIleuHISerIleuValIleValIleValIleValIle 753
DB 1612 AATGAGAACTTAAGAACTCTTGAAGATTTGAAATGTGAGAGAAAGAAAGCAT 1671
QY 753 uValIleProGlyArgGlnIleArgIleuGlnIleSerIleuSerIleuSerGlyIle 772
DB 1672 TGTAAACCTTCAATGATTAATAAGAAAGACAGATCTCAAACTACTGCTCGGAA 1731
QY 773 ---GlyLeuGlnIleuArgPhePheLeuPheAspValIleuLeuTy 790
DB 1732 CACTTACGACAAAGAGCTTACTTATTCACCAATGATGTGCTGATCGTGTGCC 1791
QY 790 rArgGlyLeuThrAlaSerAsnGlnPheIleValIleGlyIleuProLeuTyTrpGly 810

[illegible]

PRIOR FILING DATE: 2000-10-03			
PRIOR APPLICATION NUMBER: 60/231,498			
PRIOR FILING DATE: 2000-09-08			
NUMBER OF SEQ ID NOS: 207012			
SOFTWARE: FastSeq for Windows Version 4.0			
SEQ ID NO 5747			
LENGTH: 4272			
TYPE: DNA			
ORGANISM: Human			
US-09-949-016-5747			
Alignment Scores:			
Pred. No.:	3,44e-31	Length:	4272
Score:	443.00	Matches:	180
Percent Similarity:	39.1%	Conservative:	111
Best Local Similarity:	24.2%	Mismatches:	302
Query Match:	8.1%	Indels:	153
DB:	3	Gaps:	22
US-09-555-342B-2 (1-1045) x US-09-949-016-5747 (1-4272)			
QY	401	AlaIaIuSerProGluGlyGlnSerCyValArgAlaGlyLysGluProLysValSerAlaGly	420
DB	1438	GCCAGGATCCGCCAGACCCAGTCAATG-----TGCACAGC	1473
QY	421	GluProGlySerHisProSerProAlaPro-----	430
DB	1474	CAACCTGCAAGCCACCAAGTCCCAAGTCCCCCAAGGTAAGGCTCCCGCTGCTTTC	1533
QY	431	-----ArgArgSerProAlaGlyAlaLysValaSer	441
DB	1534	TGCTGGCTCTGGGCCCCGGGAGCGGTAGAAGGTGCGCAACGGGACAGCGGCAATTGATA	1593
QY	442	GlyAlaAlaSerAlaProThrGlnGluGlnGluAluValLysAspArgThrGlnGln	461
DB	1594	GCAATCACTGCGCATCCCAACAGCGAGG-----	1620
QY	462	SerLysProGlnProProGlnProSerThrGlySerLeuThrGlySerProHisLeuSer	481
DB	1621	-----AAACCTGCTTCGTCAGTATGACGGGCCCCCA-----	1653
QY	482	GluLeuSerValaIaSerGlnGlyLysValaAlaProAlaAsnValThrLeuSerProAsn	501
DB	1654	-----GCCACAGGCTCTGCCCTGGGCGCCCTGCCCTGGTAGTGCTCTGTTGCTTGG	1706
QY	502	LeuSerProAspThrLysGlnAlaSerProLeuLeuSerProLeuLeuAsnArgGlnAla	521
DB	1707	GCCGACCCGCCAGCGCTGCTCCCAAGAGTATGACATGACCTCGAGAGAGAGAC--	1763
QY	522	CysProArgThrArgAspArgLysAspGlnGlyArgArgLysArgPheProThr-----	538
DB	1764	-----GACGAGAGAGAGAGAGAGAGAGAGAAATCCCAATGCCATGCCCTGATG	1814
QY	539	-----AspLysAlaTyrrPheIleAlaLysGlu	547
DB	1815	GAGAGACAGAGATCTGTGAGATTGACTGTGCACAGAAAGAGGTTCACATGGCCAAATGAG	1874
QY	548	ValSerThrThrGlnArgLysArgLysValaSerPheHisThrAsnPhe--LeuLysGluIle	567
DB	1875	CTCTGCAAACTTGAGAGGCGCTTCCAGGCTCCATCTCCCTCGAATCAAGGTGTCTGT	1934
QY	568	SerThrValSerLysGlu-----AspAlaMetProGluAlaLeuLysSerLeu	583
DB	1935	GCCGCGGCTGCGTGAAGAAAGCTCGGAAACCGGAGTTCCTCCGCGCGGACGTTGTCCACGCG	1994
QY	584	IlePheProAsnPheGluProLeuHisLysPheHisThrAsnPhe--LeuLysGluIle	602
DB	1995	ATCTTCTCTAAACATCTGCTCATCTATTGCTTCCACAGACAGTCTCTGCTGCTCAAGCTA	2054
QY	603	GlnGlnArgLeuAlaLeuThrGlnGlyArgSerAsnAlaGlnIleArgAspTyrrGlnArg	622
DB	2055	GAGAAAGCGCATGAGAGAAATGGAC-----CGGTATTCACGC	2090


```

Db      366 TTCTCCACCTGGCTGAAACTCAATAGAAAGTAGTCCAGAGATGGCGGAAAGAAAGC 425
Qy      110 HtValValVallyPheValVallyPhePheProPhePhe--HtHrGlnLeuGln 128
Db      426 CCCCTGCTTTTAAAGTTCCGTGCCAGAGTTACCTCAGATGTGTCCGGAATGATTT 485
Qy      129 GlnGlnLeuHrPheArgTrLeuPheAlaLeuGlnVallyPheGlnPheAlaGlnGlyArg 148
Db      486 CAGGACATCACTACGCGCTGTGTCTTTCTTCGATGATGAAAGAGGCAATTCATGATGAT 545
Qy      149 LeuThrCysAsnAspThrSerAlaAlaLeuLeuLysSerHisLeValGlnSerGlnLe 168
Db      546 ATTACTACCGCCGCTGAGACCGCTGTGTCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 605
Qy      169 GlyAspPheArgGlnAlaLeuAspArgGln--HisLeuAlaLysAsnLysTrpTrLePro 187
Db      606 GGGGACTTCAATAGAAAGATGCATMAAGTCTGGCTACCTGGCCGAGAGCAAGTTGCTCCG 665
Qy      188 Gln-----GlnAspAlaLeuGlnAspLysLeValGln 198
Db      666 CAGAGACTCCTGGAACGCAACAACTCAACAGAGACGAGGAGAGCGGATCCAGGTG 725
Qy      199 PheHisHisAsnHisLeGlyGlnThrProAlaGlnSerAspPheGlnLeuLeuGlnLe 218
Db      726 TGGCATGAGAAACACCGTGGCATGTCTCAGGAGAGATGCTGTCTGTGAAATATCTGAAGAT 785
Qy      219 AlaArgArgLeuGlnMetTrpGlyTrLeArgLeuHisProAlaLysAspArgGlnGlyThr 238
Db      786 GCTCAATCTGTGAGATGTGTGTGTGAACCTTCACTGACATCAAGAAACAAGAAAGCTCA 845
Qy      239 LysLeuLeuLeuAlaValAlaAsnThrGlyTrLeLeuValPheGlnGlyPheThrLysLe 258
Db      846 GAGCTGTGCTGGGGGTGGATGCTGCTGGCTCTCAACATCTHATGAGCAAGATGACAGCTA 905
Qy      259 Asn-----AlaPheAsnTrpAlaLysValArgLysLeuSerPheLysArgLysArg 275
Db      906 ACTCCCAAGATAGGCTTCCCTGAGATGAAATGAGAAACATCTCTTCAATGATTAAGAAA 965
Qy      276 PheLeuLe-----LysLeuArgProAspAlaAsnSerAlaTrpGlnAspThr 291
Db      966 TTGTCTCTCAAGCCCATTTGACAAAAAAGCCCGGAC-----1001
Qy      292 LeuGlnPheLeuMetAlaSerArgAspPheCysLysSerPheTrpLysLeCysValGln 311
Db      1002 TTGCTCTTATGTCTCCCGGCTGCGGATTAACAAGGAGATCTTGGCTTGTGATGGGG 1061
Qy      312 HisHisAlaPhePheArgLysLeuPheGlnTrpLysProLysProValLeuPhe 331
Db      1062 AACCATGAATATACATG-----CGCGGTGCAAGCCTGATACCATTTAGAGGTG 1109
Qy      332 SerArgLysSerSerPheArgPheSerCysArgHrGlnLysGlnValLeuAspTrpVal 351
Db      1110 CAGAGATGTAAGGACAGCGCCGGAGAGAGAACCAAGAGAGAGAGAGAGCTGTATG 1169
Qy      352 LysGlnGlyGlyHisLysLysValGlnPheGlnLysTrpHisSerLys-----367
Db      1170 CTGAAATAAGAAAGAAAGAGCGTGAATGCGAAGAAAGAAAGAAAGATTGAACGG 1229
Qy      368 -----HisSerLeArgSerLeuAlaSerGlnProThrGlnLeuAsn 382
Db      1230 GAGAAAGAGAGCTGATGAGAGAGCTGAGACATGAGAAACAGATTAAGAAAGCTCAG 1289
Qy      383 SerGlnValLeuGlnSerGlnSerGlnSerThrSerLeuThrPheGly--GlnGlyAla 401
Db      1290 CAAGAACTGAAAGACAGACCGGTAGAGCTCTGAACTTGTAGCA--GGAACGGAAGCGGTG 1348
Qy      402 GlnSerProGly-----GlyGlnSerCysArgArgGly-----LysGlnProLys 416
Db      1349 CCAAGAGAGAGCTGAAAGAGCTGCGCAAGAGAGCTCAAGAGAGCTAAAGAGCGCAAGAGCG 1408
Qy      417 ValSerAlaGlyLysGlnProGlySerHisProSerProAlaProArgArgProAlaGly 436

```

```

Db      1409 CTGTGTCAGAGGCTCCCGGGA-----CCAGAAAAAGCTCAGAAACAGTGGCTTGA 1462
Qy      437 AsnLysGlnAlaAspGly-----AlaAlaSerAlaProThrGlnGln--GlnGln 452
Db      1463 AATGGCAGAGCTGACAGCTGAAATCTCCAGCTGAGAGATGGCCGACAGAAAGAGAGAG 1522
Qy      452 uGlnValVal-----LysAspArgTrpGlnGlnSerLysProGlnProProGln 468
Db      1523 TAGAGCTGTGAGATGCGACAGAGAGCCCAAGATGGTATGACAGAAAGCTTGAGAGAGCCG 1582
Qy      468 nProSerThrGlySerLeuThrGlySerProHisLysSerGlnLysSerValAsnSerGln 488
Db      1583 TGTGAGCTGAAGACCTGCATGATGACACTATGTGGAGAGAGCTTGTGAGATATAGA 1642
Qy      488 nGlyGlyValAlaProAlaAsnValThrLeuSerProLeuLeuSerProAspThrLysGln 508
Db      1643 G-----1643
Qy      508 nAlaSerProLeuLeuSerProLeuLeuAsnGlnAlaCysProAlaArgTrpAspArgGln 528
Db      1644 -----GATGAGCA 1651
Qy      528 uAspGlnGlyArgArgLysArgPheProThrAspLysValaTrpPheLeuAlaLysGlnVal 548
Db      1652 GATATGAATATGGGGCAGAGCTATGTGCTGACCTACGCGGTGATGTCTATGCGCAAGAGCG 1711
Qy      548 LserThrThrGlnArgThr-----TrpLeuLysAspLeuGln 560
Db      1712 CAGTGAAGAGAAACGTACCACTGAGGACAGAGAAAGATAGCTGTGACAGAGACCTGAA 1771
Qy      560 uValLeuThrSerTrpPheGlnSerThrValSerLysGlnAspAlaMetProGlnAlaLe 580
Db      1772 GGGCTCACTTG-----GAGCTGAGCAATGCGAGATGAGTCCAGAAAGAGCTGC 1822
Qy      580 uLysSerLeuLysPheProAsnPheGlnProLeuHisLysPheHisThrAsnPheLeu 600
Db      1823 CAATGACATGATCCATGCTGAGAAACATGAGACTGGGCGAGACCAATTAACAGACCTTGG 1882
Qy      600 sGlnLeuGlnArgLysLeuAlaLeuTrpGlnGlyArgSerAsnAlaGlnLeuArgAspTrp 620
Db      1883 CCAAGTCCGCGAG-----GGCAACACCAAGACCGCATTTAGAGAAAT 1924
Qy      620 rGlnArgLys-----GlyAspValMetLeuLysAsnLeuGlnGln 633
Db      1925 TAGGTATATGATATGGGACCCAGACCTCTAGGAGACCCCTCCCTTTTCCCTTGTCCCC 1984
Qy      633 yMetLysHisLysLeuAlaHisLeuTrpLysHisSerGlnAlaLeuGlnAlaLeuGlnLys 653
Db      1985 ACACTCTTACAC--CTAATCTACCTA--ACTCATCT--GTGCTGAGGACCTAATCTAG 2037
Qy      653 nGlyLysSerSerArgArgLeuGlnAsnPheCysArgAspPheGlnLeuGlnLys--672
Db      2038 AGCAGCCCTGAGATCA-----TCCCAAGCATTTAAATGATGCCATGGG 2079
Qy      672 -----672
Db      2080 ACCAAACTAGCCCTTAGCCCCCAACCACTTCCCTGGCAATATGATGGCTCATATGG 2139
Qy      673 -----ValCysTrpLeuProLeuAsnThrPhe-----681
Db      2140 TGGCAATGAAACCTCTTCTCTTCTGTGTCTGTATGAATCTGTATGAGGCTGAGAAATCTCT 2199
Qy      682 -----LeuLeuArgProLeuHisArgLeuMet 690
Db      2200 ACTTCTCAAGCTTAAGATCTTTCATCTGATTTTGGCAAAATCCCTTACACTTACTGT 2259
Qy      690 HisTrp--LysGlnValLeuGlnArgLysCysLysHisHisProPro 705
Db      2260 GTCTATGGGAGTCAAGTGTGAGTAGTGGAGTGGAGACTGCTCCCTCT 2308

```

Search completed: May 20, 2006, 03:27:44
Job time : 548 secs

This Page Blank (uspto)

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: May 20, 2006, 03:21:43 ; Search time 2764 Seconds
(without alignments)
6968.476 Million cell updates/sec

Title: US-09-555-342B-2

Perfect score: 5463
Sequence: 1 MGIEBGRPRFGSLGAPENS.....SATSSARPHVLSHKSLEY 1045

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 18892170 seqs, 6143817638 residues

Total number of hits satisfying chosen parameters: 3784340

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing filter 45 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=xlh
-Qc/abgs/abssweb.spool/US09555342/runcat.19052006.121509.13438/app_query.fasta_1
-DB=Published Applcations NA Main -QPMT=fastcap -SUFFIX=p2n.rnpbm
-MINMATCH=0.1 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1
-MATRIX=blomsu62 -TRANS=human0.cdt -LIST=45 -DOCALIGN=200 -THR.SCORE=pct
-THR.MX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL -OUTPMT=apco -NORM=ext
-HEAPSIZE=500 -MTLEN=0 -MAXLEN=200000000 -HOST=abs02h
-USER=US09555342 @CCG 1.1.1675 @runcat.19052006.121509.13438 -NCPU=6 -ICPU=3
-NO MAMP -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-MAMP TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPO=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications NA Main:
1: /EMC Celerra SIDS3/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /EMC Celerra SIDS3/ptodata/2/pubpna/US08_PUBCOMB.seq:*
3: /EMC Celerra SIDS3/ptodata/2/pubpna/US09_PUBCOMB.seq:*
4: /EMC Celerra SIDS3/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
5: /EMC Celerra SIDS3/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
6: /EMC Celerra SIDS3/ptodata/2/pubpna/US10_PUBCOMB.seq:*
7: /EMC Celerra SIDS3/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
8: /EMC Celerra SIDS3/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
9: /EMC Celerra SIDS3/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
10: /EMC Celerra SIDS3/ptodata/2/pubpna/US10F_PUBCOMB.seq:*
11: /EMC Celerra SIDS3/ptodata/2/pubpna/US10G_PUBCOMB.seq:*
12: /EMC Celerra SIDS3/ptodata/2/pubpna/US10H_PUBCOMB.seq:*
13: /EMC Celerra SIDS3/ptodata/2/pubpna/US11A_PUBCOMB.seq:*
14: /EMC Celerra SIDS3/ptodata/2/pubpna/US11B_PUBCOMB.seq:*
15: /EMC Celerra SIDS3/ptodata/2/pubpna/US11C_PUBCOMB.seq:*
16: /EMC Celerra SIDS3/ptodata/2/pubpna/US11D_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5463	100.0	3442	7	US-10-159-563-122

2	5459	99.9	4687	9	US-10-921-707-18	Sequence 18, App1
3	5431	99.4	3187	7 <th>US-10-450-763-554</th> <th>Sequence 564, App</th>	US-10-450-763-554	Sequence 564, App
4	2941.5	53.8	3997	10 <th>US-10-172-118-1536</th> <th>Sequence 1536, App</th>	US-10-172-118-1536	Sequence 1536, App
5	2941.5	53.8	3997	8 <th>US-10-342-887-1536</th> <th>Sequence 1536, App</th>	US-10-342-887-1536	Sequence 1536, App
6	2941.5	53.8	3997	3 <th>US-10-802-432-16</th> <th>Sequence 16, App1</th>	US-10-802-432-16	Sequence 16, App1
7	2498	45.7	3094	3 <th>US-09-925-297-220</th> <th>Sequence 220, App</th>	US-09-925-297-220	Sequence 220, App
8	1006.5	18.4	1718	3 <th>US-09-764-868-51</th> <th>Sequence 51, App1</th>	US-09-764-868-51	Sequence 51, App1
9	913.5	16.7	1747	13 <th>US-11-097-143-3041</th> <th>Sequence 3041, App</th>	US-11-097-143-3041	Sequence 3041, App
10	905.5	16.6	1220	13 <th>US-11-097-143-743</th> <th>Sequence 743, App</th>	US-11-097-143-743	Sequence 743, App
11	891.5	16.3	4150	13 <th>US-11-097-143-3040</th> <th>Sequence 3040, App</th>	US-11-097-143-3040	Sequence 3040, App
12	748.5	13.7	3309	8 <th>US-10-755-889-499</th> <th>Sequence 499, App</th>	US-10-755-889-499	Sequence 499, App
13	748.5	13.7	3309	10 <th>US-10-956-157-1800</th> <th>Sequence 1800, App</th>	US-10-956-157-1800	Sequence 1800, App
14	744	13.6	4543	8 <th>US-10-152-119A-2048</th> <th>Sequence 2048, App</th>	US-10-152-119A-2048	Sequence 2048, App
15	744	13.6	4543	16 <th>US-11-036-196-2048</th> <th>Sequence 2048, App</th>	US-11-036-196-2048	Sequence 2048, App
16	743.5	13.6	3387	8 <th>US-10-755-889-397</th> <th>Sequence 397, App</th>	US-10-755-889-397	Sequence 397, App
17	735.5	13.5	14917	3 <th>US-09-909-567B-11</th> <th>Sequence 11, App1</th>	US-09-909-567B-11	Sequence 11, App1
18	724	13.3	3620	6 <th>US-10-084-817-280</th> <th>Sequence 280, App</th>	US-10-084-817-280	Sequence 280, App
19	724	13.3	3620	9 <th>US-10-921-707-17</th> <th>Sequence 17, App1</th>	US-10-921-707-17	Sequence 17, App1
20	724	13.3	4336	6 <th>US-10-171-581-5</th> <th>Sequence 5, App1</th>	US-10-171-581-5	Sequence 5, App1
21	724	13.3	4336	7 <th>US-10-159-563-406</th> <th>Sequence 406, App</th>	US-10-159-563-406	Sequence 406, App
22	713	13.1	4922	9 <th>US-10-357-930-30232</th> <th>Sequence 30232, App</th>	US-10-357-930-30232	Sequence 30232, App
23	712.5	13.0	6263	10 <th>US-10-756-149-4222</th> <th>Sequence 4222, App</th>	US-10-756-149-4222	Sequence 4222, App
24	712.5	13.0	6263	10 <th>US-10-765-700-57</th> <th>Sequence 57, App1</th>	US-10-765-700-57	Sequence 57, App1
25	703	12.9	3643	9 <th>US-10-772-636-39</th> <th>Sequence 39, App1</th>	US-10-772-636-39	Sequence 39, App1
26	703	12.9	3643	16 <th>US-11-288-493-39</th> <th>Sequence 39, App1</th>	US-11-288-493-39	Sequence 39, App1
27	700	12.8	2882	10 <th>US-10-764-420-2568</th> <th>Sequence 2568, App</th>	US-10-764-420-2568	Sequence 2568, App
28	699.5	12.8	5897	13 <th>US-11-097-143-24557</th> <th>Sequence 24557, App</th>	US-11-097-143-24557	Sequence 24557, App
29	695.5	12.7	4125	6 <th>US-10-044-090-620</th> <th>Sequence 620, App</th>	US-10-044-090-620	Sequence 620, App
30	691.5	12.7	7982	13 <th>US-11-097-143-742</th> <th>Sequence 742, App</th>	US-11-097-143-742	Sequence 742, App
31	690.5	12.6	2916	13 <th>US-11-097-143-15944</th> <th>Sequence 15944, App</th>	US-11-097-143-15944	Sequence 15944, App
32	685.5	12.5	3470	13 <th>US-11-097-143-24407</th> <th>Sequence 24407, App</th>	US-11-097-143-24407	Sequence 24407, App
33	675	12.4	393	3 <th>US-09-918-995-8794</th> <th>Sequence 8794, App</th>	US-09-918-995-8794	Sequence 8794, App
34	639	11.7	506	3 <th>US-09-967-768A-104</th> <th>Sequence 104, App</th>	US-09-967-768A-104	Sequence 104, App
35	639	11.7	506	10 <th>US-10-843-614A-6249</th> <th>Sequence 6249, App</th>	US-10-843-614A-6249	Sequence 6249, App
36	626	11.5	3166	9 <th>US-10-370-715B-537</th> <th>Sequence 537, App</th>	US-10-370-715B-537	Sequence 537, App
37	626	11.5	3398	8 <th>US-10-302-172-201</th> <th>Sequence 201, App</th>	US-10-302-172-201	Sequence 201, App
38	609.5	11.2	716	3 <th>US-09-764-868-475</th> <th>Sequence 475, App</th>	US-09-764-868-475	Sequence 475, App
39	608.5	11.1	2067	8 <th>US-10-250-613-23</th> <th>Sequence 23, App1</th>	US-10-250-613-23	Sequence 23, App1
40	605	11.1	3121	7 <th>US-10-108-260A-1532</th> <th>Sequence 1532, App</th>	US-10-108-260A-1532	Sequence 1532, App
41	596.5	10.9	717	7 <th>US-10-264-237-1390</th> <th>Sequence 1390, App</th>	US-10-264-237-1390	Sequence 1390, App
42	590.5	10.8	379	3 <th>US-09-960-352-154</th> <th>Sequence 154, App</th>	US-09-960-352-154	Sequence 154, App
43	590	10.7	2846	8 <th>US-10-415-187-17</th> <th>Sequence 17, App1</th>	US-10-415-187-17	Sequence 17, App1
44	587	10.7	5582	15 <th>US-11-000-688-386</th> <th>Sequence 386, App</th>	US-11-000-688-386	Sequence 386, App
45	586.5	10.7	4356	13 <th>US-11-097-143-42194</th> <th>Sequence 42194, App</th>	US-11-097-143-42194	Sequence 42194, App

ALIGNMENTS

RESULT 1
US-10-159-563-122
Sequence 122, Application US/10159563
Publication No. US20040009154A1
GENERAL INFORMATION:
APPLICANT: Khan, Javed
APPLICANT: Peterson, Markus
APPLICANT: Mettzer, Paul
TITLE OF INVENTION: SELECTIONS OF GENES AND METHODS OF USING THE SAME FOR
FILE REFERENCE: 11613.56US11
CURRENT APPLICATION NUMBER: US/10/159,563
CURRENT FILING DATE: 2002-12-09
PRIORITY APPLICATION NUMBER: US 10/133,937
PRIORITY FILING DATE: 2002-04-25
NUMBER OF SEQ ID NOS: 444
SOFTWARE: PatentIn version 3.1
SEQ ID NO 122
LENGTH: 3442
TYPE: DNA
ORGANISM: Homo sapiens
US-10-159-563-122
Alignment Scores:

Pred. No.: 0 Length: 3442
Score: 5463.00 Matches: 1045
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 7 Gaps: 0

US-09-555-342B-2 (1-1045) x US-10-159-563-122 (1-3442)

QY 1 MetGlyGluIleGluGlnAaGProThrProGlySerArgLeuGlyAlaProGluAaSer 20
DB 49 ATGGAGAAATTAAGCAGAGCCGACCCAGAGATCAGACTGGGGCCCGGAAATTCG 108
QY 21 GlyIleSerThrLeuGluAaGlyGlnLysProProProThrProSerGlyLysLeuVal 40
DB 109 GGGATCATGTACCTTGGAACTGGACAGAAAGCCGCCCAACCTTCAGAAAATCTGTG 168
QY 41 SerIleLysIleGlnMetLeuAaPsrThrGlnGluAlaPheGluValProGlnArgAla 60
DB 169 TCCATCAAAATCCAGATGCTGGATGACACCAGAGGCACTTGAAGTTCCAAAGAGCT 228
QY 61 ProGlyLysValLeuLeuAaPAlaValCysAaAhiLysLeuAaLysValGluGlyAaPTr 80
DB 229 CCGGGAGAGGTGCTGTGAGATGACAGTTTGCAACCACTCAACTCGTGGAAAGGTGACTAT 288
QY 81 PheGlyLeuGluPheProAaPAlaLysLysIleThrValTrpLeuAaPLeuLysPro 100
DB 289 TTTGGCTCGAGTTTCTGTGATCAAAATAATCAGGTGGCTGGATCTCTTAAACCC 348
QY 101 IleValLysGlnIleArgArgProLysIleValIleLysPheValIleLysPhePhe 120
DB 349 ATTGTGAACAGATTAGAGGCCAAAGCAGTGTGTTAAGTTGTGGTGAATTTCTT 408
QY 121 ProProAaPAlaThrGlnLeuGlnGluLeuThrArgTrpLeuPheAlaLeuGlnVal 140
DB 409 CCCCCTGACCAACACAACTCCAGAGAACTCAAGAGTACCTGTTCGGCTGCAGGTG 468
QY 141 LysGluAaPLeuLysGlnGlyArgLeuThrCysAaAaPThrSerAlaAlaLeuLysIle 160
DB 469 AAGCAGACCTTGGCTCAAGGCAAGTTGACGTGTATACACAGCGAGCTCTTGATT 528
QY 161 SerIleIleValGlnSerGluIleGlyAaPheAaPAlaLeuAaPArgGluIleLeu 180
DB 529 TCACACATTTGTGCATATGAGATTGGGGATTTGTGATGAACCTTGGACAGAGCACTTA 588
QY 181 AlaLysAaLysTrpIleProGlnGlnAaPAlaLeuGluAaPLeuValIleValGluPheHis 200
DB 589 GCAAAAATAATATACATACCTCAGCAAGACGCACTAGAGAGCAAAATCGTGAATTTGCAC 648
QY 201 HisAaAhiLysIleGlnThrProAlaGluSerAaPheGlnLeuLeuGluIleAlaArg 220
DB 649 CATACACACATTTGGACAAACACAGCGAATTCAGATTTCCAGCTCCAGAGATTTGCCCGT 708
QY 221 ArgLeuGlnMetTrpGlyIleArgLeuHisProAlaLysAaPArgGluIleTrpLysIle 240
DB 709 CGGCTAGAGATGTATGGAAATCCGGTTGCACCCGGCCAAAGACAGGGAGGCAAGAGATC 768
QY 241 AsnLeuAlaValAlaAaPThrGlyIleLeuValPheGlnGlyPheTrpLysIleAaAla 260
DB 769 AATCTGGCTCTGGCAACAGGGAATCTAGTGTTCAGGGGTTTCACTAAGATCAAGGCC 828
QY 261 PheAaPTrpAlaLysValArgLysLeuSerPheLysArgLysArgPheLeuIleLysLeu 280
DB 829 TTTCACTAGGGCCAAAGGTCGGAAGCTGACTTCAGAGAGAAAGGCTTTCTCAACAAGCTC 888
QY 281 ArgProAaPAlaAaSerAlaLysGlnAaPThrLeuGlnPheLeuMetAlaSerArgAaP 300
DB 889 CGGCGAGATGCAATATGTGCTACAGAGATACCTTGAATTTCCGATGGCCAGTGGGAT 948
QY 301 PheCysLysSerPheTrpLysIleCysValGluHisIleAlaPhePheAaGluPheGlu 320
DB 949 TTCTGCAAGTCTTTGGAAAATCTGTGTGAACATCATGCTTTTGAACCTTTTGA 1008

QY 321 GluProLysProLysProLysProValLeuPheSerArgGlySerSerPheArgPheSer 340
DB 1009 GAGCCCAAAACCAAGGCCAAGCCCTCTCTTTAGCCGGGGGTCACTATTCGGTTCAGT 1068
QY 341 GlyArgThrGlnLysGlnValLeuAaPTrpValLysGluGlyValLysLysValGln 360
DB 1069 GGTGCGACTCAGAAAGCAGGTTCTCGACTATGTAAAGAAAGAGACATAAAGAGGTGCAG 1128
QY 361 PheGluAaPLeuHisSerLysIleHisSerIleArgSerLeuAlaSerGlnProGlnGlu 380
DB 1129 TTTGAAAGAGACACAGAAATTCATCTATCCAGGCTTGCTTCAAGCTTACAGAA 1188
QY 381 LeuAaPSerGlnValLeuGlnGlnSerGlnGlnSerThrSerLeuThrPheGlyGluGly 400
DB 1189 CTGAATTCGGAAGTGTGGAGAGAGCTTCAGAGAGACACAGCTTTCATTTGGAAAGT 1248
QY 401 AlaGluSerProGlyGlyGlnSerCysArgArgGlyLysGluProLysValSerAlaGly 420
DB 1249 GCCGAATCTCAGGGGGCAGAGCTCCGGCGAGAAAGAACCGAAAGTTTCCGGCCGG 1308
QY 421 GluProGlySerHisProSerProAlaProArgArgSerProAlaLysLysGlnAla 440
DB 1309 GAGCCGGGGTCCGACCCGAGCCCTGCGCGAGAGAAAGCCCGGGGTAAACAAGCGG 1368
QY 441 AspGlyAlaAlaSerAlaProThrGlnGluGluGluValIleLysAaPArgTrpGln 460
DB 1369 GACGAGACCGGCTCGGCGCCACCGAGAAAGAGAGAGAGGTGTTAAGGATAGACCCAG 1428
QY 461 GlnSerLysProGlnProProGlnProSerThrGlySerLeuThrGlySerProHisLeu 480
DB 1429 CAGAGTAAACTCTACGCCCCCGCAGCCACACAGGCTCTGACTGCGAGTCTCACTT 1488
QY 481 SerGluLeuSerValAaPserGlnGlyValAlaProAlaAaPAlaThrLeuAaPPro 500
DB 1489 TCCAGCTGTCTGTGAATCTGAGGGGGAGAGGCCCCCTGCAAGCTTATGTCTCC 1548
QY 501 AsnLeuSerProAaPThrLysGlnAlaSerProLeuIleSerProLeuAaPLeuAaPArg 520
DB 1549 AACCTAGGCCCGGACACCAAGCAGGCTCTCTCTTATCAGCCGCTGGATGACAG 1608
QY 521 AlaCysProArgThrAaPArgLysAaPArgGluGlyArgArgLysPheProThrAaPArg 540
DB 1609 GCCTGCCCCCGGACGACATGAGTAGGGCCCGGAGAAAGATTTCCACATGTATAA 1668
QY 541 AlaTrpPheIleAlaLysGlnValSerThrTrpGlnAaPThrLysLeuLysAaPLeuGlu 560
DB 1669 GCGTACTTCACTAGAGAGAGTGTACCAACCGAGCAACATATGTGAAGATCTCGAA 1728
QY 561 ValIleThrSerTrpPheGlnSerThrValSerLysGluAaPAlaMetProGluAlaLeu 580
DB 1729 GTTATCACTTCGTCGTTTCAAGACAGTAGAGCAAGAGAAAGAGACGCAATGCCGAAAGCACTG 1788
QY 581 LysSerLeuIlePheProAaPheGluProLeuHisLysArgPheHisThrArgPheLeuLys 600
DB 1789 AAAAGTCAATATTCGGAATTTGAACTTTGCAAAATTTCAATCTTAATTTTCCAAAG 1848
QY 601 GluIleGlnGlnArgLeuAlaLeuTrpGluGlyArgSerAaPAlaGlnIleArgAaPTr 620
DB 1849 GAATATTGACAAAGACTTGCTGCTGGAGAGCCGCTCAATATCCCAATCAAGATTTAC 1908
QY 621 GlnArgIleGlyAaPValMetLeuLysAaPLeuIleGlnGlyMetLysHisLeuAlaAlaHis 640
DB 1909 CAAGAGATCGGAGATGTCACTGGAAGAACTTCAAGGCACTGAAGCACTGGCGGCTCAC 1968
QY 641 LeuTrpLysHisSerGlnAlaLeuGluAlaLeuGluAaPArgGlyLysSerSerArgArg 660
DB 1969 CTGTGAAACACACGACGAGGCTTTGAGGCTCTGGAATAAGATTCAGAGCTCCGGCGG 2028
QY 661 LeuGluAaPheCysArgAaPheGluLeuGlnLysValCysTrpTrpProLeuAaPThr 680
DB 2029 CTGGAAATCTTCTGCAAGATCTTGAAGCTGCGAAGAGTGTGTTAATCCAGCTCAACACC 2088
QY 681 PheLeuLeuArgProLeuHisArgArgLeuMetHisTrpLysGlnValLeuGluAaPLeuCys 700

Db 2089 TTCCTCGGCGCCAGCTGACCGGCTCATGACATGACAGAGCTCTGAGCGGCTGCG 2148
Qy 701 LysHSHSPROFROSERHIALAAPHARAPCYSAAGLAAALAEUALEGLUTLE 720
Db 2149 AACAACACCCCGGAGCCAGCCGACCTTCAGGGAGCTCCAGCGCTTGGCAGAGATC 2208
Qy 721 ThriGmetValAlaGlnLeuHISGLYThrMetIleLysMetGluAanPheGlnLysLeu 740
Db 2209 ACCGAGATGGTGGACAGCTCCAGGTCAGATGATCAAGATGAGAAATTTCCAGAGCTG 2268
Qy 741 HisGluLeuLysLysAspLeuIleGlyIleAspAanLeuValProGlyArgGlnPhe 760
Db 2269 CACGAACCTCAAGAAAGATTGATTGAGCATTCGATGATGTTGGTTCCGGAAGGAGTTC 2328
Qy 761 IleArgLeuGlySerLeuSerLysLeuSerGlyLysGlyLeuGlnGlnArgMetPhePhe 780
Db 2329 ATCCGTCTGGGACCTCCAGCAAGCTCTCGGGGGAAGGGGCTCCAGCAGCATGTTCTTC 2388
Qy 781 LeuPheAanAspValIleLeuLysThrSerArgGlyLeuThrAlaSerArgGlnPheLys 800
Db 2389 CTGTTCAACGACGCTCTGCTATACAGCCGGGGGCTGAGCGCTCCAAATCAGTTTAAA 2448
Qy 801 ValHisGlyGlnLeuProLeuLysGlyMetThrIleGlnGlnSerGluAspGluTrpGly 820
Db 2449 GTCCAGGGGAGCTCCGGCTCTAGGATGACATTCGATGAGAGAGCAGAGAGAGTGGGG 2508
Qy 821 ValProHisCysLeuThrLeuArgGlyGlnArgGlnSerIleIleValAlaLysSer 840
Db 2509 GTGGCCCACTGGCTCGACCTCCGGGGGAGCGGAGCTCCATCATCTGGCCGCAATTCCT 2568
Qy 841 ArgSerGlnMetGluLysTrpValGluAspIleGlnMetAlaIleAspLeuAlaGlnLys 860
Db 2569 CGGTCCAGATGAGAAAGTGGGTTGAGGACATCCAGATGGCATTGACCTGGCCGAGAG 2628
Qy 861 SerSerSerProAlaProGluPheLeuAlaSerSerProProAspAsmLysSerProAsp 880
Db 2629 AGCAGCAGCCCGCCCTCGAGTCTCTGGCCAGCAGCCCGCTGACAAACATCCCTGAT 2688
Qy 881 GluAlaThrAlaAlaAspGlnGlnSerGluAspLeuSerAlaSerArgThrSerLeu 900
Db 2689 GAACCCACCGCGCTGACAGAGATCAGAGATACCTGACGCGCTCGCCGACATCCGCTG 2748
Qy 901 GluArgGlnAlaProHisArgGlyAsnThrMetValHisValCysTrpHisArgAsnThr 920
Db 2749 GAGCGCCAGGCGCCGACCGGCGAACAAATGGTGACGTGCTGCGACCGGACACCC 2808
Qy 921 SerValSerMetValaAspPheSerIleAlaValGluAanGlnLeuSerGlyAsnLeu 940
Db 2809 AGCGTCTCCATGGTGAATTGACATCGCAGATGAGAAATCAGTTGTCTGGAACCTGCTG 2868
Qy 941 ArgLysPheLysAanSerAanGlyTrpGlnLysLeuTrpValValPheThrAsnPheCys 960
Db 2869 AGGAAATTCAAAACACCAACGGGTGGAGAAAGCTGGGTGGTTCACAAACTTCTGCG 2928
Qy 961 LeuPhePheTrpLysSerHisGlnAspAsnHisProLeuAlaSerLeuProLeuLysGly 980
Db 2929 CTGTTCTTCAAAATACACACAGAGACAAATATCCCTGCGACCGCTCTCTGCGGC 2988
Qy 981 TyrSerLeuThrIleProSerGlnSerGluAanIleGlnLysAspTrpValPheLysLeu 1000
Db 2989 TACTCGCTCAACATCCCTCGAGTCCGAGAAACATCCAGAAAGCTACGTGTTCAACTG 3048
Qy 1001 HisPheLysSerHisValLysTrpPheArgAlaGlnSerGluLysTrpPheGlnArgTrp 1020
Db 3049 CACTTCAAGTCCCACTGCTACTTCAAGGCGGAAAGCGAGTTCAGAGTTGAAAGGTGG 3108
Qy 1021 MetGluValIleAspSerAlaThrSerSerAlaSerArgProHisValLysSerHisLys 1040
Db 3109 ATGGAAGTGAATCCGACGTGCCACAGCTCTGCTCGGAGCCCACTGTTGAGCCACAAA 3168
Qy 1041 GluSerLeuValLys 1045

Db 3169 GAGTCTCTTGATAT 3183
RESULT 2
US-10-921-707-18
; Sequence 18, Application US/10921707
; Publication No. US2005003447A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE PHARMACEUTICALS, INC.
; APPLICANT: LAL, Preeti
; APPLICANT: TANG, Y. Tom
; APPLICANT: YUE, Henry
; APPLICANT: HILLMAN, Jennifer L.
; APPLICANT: BANDMAN, Olga
; APPLICANT: CORLEY, Neil C.
; APPLICANT: GUGLER, Karl J.
; APPLICANT: PATTERSON, Chandra
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: BAUGHN, Mariah R.
; TITLE OF INVENTION: HUMAN CYTOSKELETON ASSOCIATED PROTEINS
; FILE REFERENCE: PF-0594 PCT
; CURRENT APPLICATION NUMBER: US/10/921,707
; CURRENT FILING DATE: 2004-08-19
; PRIOR APPLICATION NUMBER: US/09/786,797
; PRIOR FILING DATE: 2001-06-25
; PRIOR APPLICATION NUMBER: 09/156,470; unassigned: 60/131,321
; PRIOR FILING DATE: 1998-09-18; 1998-09-18; 1999-04-27
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PERL Program
; SEQ ID NO 18
; LENGTH: 4687
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 1320252
US-10-921-707-18
Alignment Scores:
Pred. No.: 0 Length: 4687
Score: 5459.00 Matches: 1044
Percent Similarity: 99.9% Conservative: 0
Best Local Similarity: 99.9% Mismatches: 1
Query Match: 99.9% Indels: 0
Gaps: 0
US-09-555-342B-2 (1-1045) x US-10-921-707-18 (1-4687)
Qy 1 MetGlyGluIleGlnGlnArgProThrProGlySerArgLeuGlyAlaProGluAanSer 20
Db 90 ATGGAGAAATTAAGACAGAGCGCCAGCCAGATCAAGACTGGGGGCCCCGGAATAATTCG 149
Qy 21 GlyIleSerThrLeuGlnArgGlyGlnLysProProProThrProSerGlyLysLeuVal 40
Db 150 GGGATCAGTACCTTGGAACTGGACAGAAACCCGCCCAACACTTCAGGAAAACTGCTG 209
Qy 41 SerIleLysIleGlnMetLeuAspAspThrGlnGluAlaPheGlnValProGlnArgAla 60
Db 210 TCAATCAAAATCCAGATGCTGAGATGACACCCAGAGAGCAATTTGAAGTTCCACAAAGAGCT 269
Qy 61 ProGlyLysValIleuLeuAspAlaValCysAsnHisLeuAanLeuValGluGlyAspTrp 80
Db 270 CTGGGAGAGGTGCTGCTGAGATGACAGTTTGCACACCACTCAACCTCGGGAAGGTGACTAT 329
Qy 81 PheGlyLeuGlnPheProAspHisLysLysLysIleThrValTrpLeuAspLeuLysPro 100
Db 330 TTTGGCTCTCAGATTCTCTGATCACAAAAGATACGGTGGGTGGATCTCTCAAAAACCC 389
Qy 101 IleValIleGlnIleArgArgProLysHisValValLysPheValValLysPhePhe 120
Db 390 ATTGGAACAGATTGAAGGCCAAGACACAGTGTGTAAAGTTGTGTGTAATTTCTTT 449
Qy 121 ProProAspHisArgGlnLeuGlnGlnGluLeuThrArgTrpLeuPheAlaLeuGlnVal 140

Db 450 CCGCTGACCACACAACTCCAAAGAACTCAAGGTAACCTGTTCCGGCTGCAGAGT 509
Qy 141 LYGGLAAspLeuL1aGInG1YArgLeuThrCysAsnAspThrSerL1aL1aLeuL1e 160
Db 510 AACGAGACTGGCTCAAGGACAGGTGACGTGAATGACACAGCGAGCTCTTGATT 569
Qy 161 SerH1eValGInserGluI1eG1YAspPheAspGluL1aLeuAspArgGluH1eL1e 180
Db 570 TCACACTTGTGCATCTGAGATTGGGGATTGATGTAAGACCTTGGACAGAGCACTTA 629
Qy 181 ALALysenLysTr1eProGInGInAspL1aLeuG1uAspLysL1eValG1uPheH1s 200
Db 630 GCAAAAAATTAATACCTCACTGACAGAACGCACTGAGAGCAAAATCGGGAATTTCCAC 689
Qy 201 H1aAsnH1s1eG1YInThrProL1aG1uSerAspPheG1uInLeuG1uI1eAl1aArg 220
Db 690 CATTAACCACTTGGACAAACACAGCAAGATCCAGATTTCCAGCTCTTAGGATGGCCGT 749
Qy 221 ArgLeuG1uMetTrG1Y1eArgLeuH1sProL1aLysAspArgG1uG1YThrLys1e 240
Db 750 CCGCTAGAGATGATGGAATCCGGTTGCAACCGGCCAAGGACAGGGAGGCAAGAGATC 809
Qy 241 AsnLeuL1aValL1aAsnThrG1Y1eLeuValPheGInG1YpHeThrLys1eAsnL1a 260
Db 810 AATCTGGCCCTGACCAACAGGGAAATCTAGTGTTCAGGGGTTTCACTAAGATCAATGCC 869
Qy 261 PheAsnThrL1aLysValArgLysLeuSerPheLysArgLysArgPheL1eLysLeu 280
Db 870 TTCACTGGGGCAAGGGCGGAAGCTGAGCTTCAAGAGGAAGCGCTTTTCAACAAGCTC 929
Qy 281 ArgProAspL1aAsnSerL1aTrG1uAspThrLeuG1uPheL1uMetL1aSerArgAsp 300
Db 930 CGGCGAGATGCCAATAGTGGTACAGGATACCTTGGAAATTCCTGATGGCAATCGGGAT 989
Qy 301 PheCysLysSerPheTrpLys1eCysValGluH1sH1aL1aPheAsnArgLeuPheG1u 320
Db 990 TTCTGCAAGCTCTTGGAAATCTGTGTGAACATCATATCCCTTTTAACTTTTGGAA 1049
Qy 321 GluProLysProLysProLysProValLeuPheSerArgLysSerPheArgPheSer 340
Db 1050 GAACCCCAACCAAGCCCAAGCCGCTCTTTAAGCGGGGGTCAATTCGATTCAGT 1109
Qy 341 G1YArgThrGInLysGInValLeuAspTrValLysG1uG1YH1sLysLysValGIn 360
Db 1110 GGTGGACTCAGAAAGAGGCTTCTGACTATGTTAAGAAGAGAGCACTAAGAAGGTGCAG 1169
Qy 361 PheGluArgLysH1sSerLysL1eH1sSerL1eArgSerLeuL1aSerGInProThrG1u 380
Db 1170 TTGGAAGGAAGCAAGCAAGATTCATTTCTATCCGGAGCCTTCTCACAGCTTACAGA 1229
Qy 381 LeuAsnSerG1uValLeuG1uGInserGInGInSerThrSerLeuThrPheG1YG1uG1Y 400
Db 1230 CTAAATTCGAAAGTGTGGAGCACTTCAACAGACCAAGCTTACATTTGAGAAAGT 1289
Qy 401 ALAG1uSerProG1YG1YGInserCysArgArgG1YLysG1uProLysValSerL1aG1Y 420
Db 1290 GCGGAATCTCAGAGGGGGCCAGAGCTCGCGGAGAAAGAACCGAAGTTTCGCGCGG 1349
Qy 421 GluProG1YSerH1sProSerProL1aProArgArgSerProL1aG1YAsnLysGInL1a 440
Db 1350 GAACCGGGGTGCGACCGGACCTGCGCCGAGAGAAAGCCCGGGGTAAACAAGCGCG 1409
Qy 441 AspG1YAL1aL1aSerL1aProThrG1uG1uG1uG1uValLysAspArgTrGIn 460
Db 1410 GAGAGAGCCCGCTCGGGGCCACGAGAGAAAGAGAGGTCTGTTAAGATAGAACCCAG 1469
Qy 461 GInserLysProGInProProGInProSerThrG1YSerLeuThrG1YSerProH1sLeu 480
Db 1470 CAAAGTAAACTCAGAGCCCGGAGCCAAAGCAAGGCTCCCTGACTGGCAATCCCACTT 1529
Qy 481 SerG1uLeuSerValAsnSerGInG1YValL1aProL1aAsnValThrLeuSerPro 500
Db 1530 TCCGAGCTGTCTGAACTCGAGGGGGAGTGGCCCTCGCAACGTGAACCTTGTCTCCC 1589

Qy 501 AsnLeuSerProAspThrLysGInL1aSerProLeuL1eSerProLeuLeuAsnArgGIn 520
Db 1590 AACCTAGCCCGGACACCAAGAGAGGCTCTCCCTTATGATAGCCGCTGATGATGACAG 1649
Qy 521 ALACysProArgTrhAspArgG1uAspG1uG1YArgArgLysArgPheProThrAspLys 540
Db 1650 GCTGGCCCCCGGACGAGATGAGTGAAGGCGCGGAGGAAGATTTCCCACTGTATAA 1709
Qy 541 ALATrPheL1aL1eG1uValSerThrThrG1uArgThrThrLeuLysAspLeuG1u 560
Db 1710 GGTACTTATAGCTAAGAGAGTGTATACCGAGCGAACAATCTTGAAGATCTCGAA 1769
Qy 561 ValL1eThrSerTrpPheGInSerThrValSerLysG1uAspL1aMetProG1uL1aLeu 580
Db 1770 GTTATCACTTGTGGTTCAGAGCAAGTGAAGAAAGAGAGCCCATGCGGAAAGCACTG 1829
Qy 581 LysSerLeuL1ePheProAsnPheG1uProLeuH1sLysPheH1sThrAsnPheLys 600
Db 1830 AAAAGTCTATATTCGGAATTTTGAACCTTTGCACAAAATTTCACTAATTTTCTCAAG 1889
Qy 601 GluL1eG1uGInArgLeuL1aLeuTrpG1uG1YArgSerAsnL1aGInL1eArgAspTr 620
Db 1890 GAATTTGACCAACGACTTCCCTGTGGAAAGCGCTCAAAATGCCAAATCAGAGATTAAC 1949
Qy 621 G1uArgL1eG1YAspValMetLeuLysAsnL1eGInG1YMetLysH1sLeuL1aL1aH1s 640
Db 1950 CAAGAATCGCGGATGTCACTGCTGAAGAACATTCAGGGCATGAAGCACTGGCGGCTCAC 2009
Qy 641 LeuTrpLysH1sSerG1uL1aLeuG1uL1aLeuG1uAsnG1YL1eLysSerSerArgArg 660
Db 2010 CTGTGGAAACAACAGCAGCGCTTGGAGCGCTCGAGAAATGGAATCAAGAGCTTCCCGCG 2069
Qy 661 LeuG1uAsnPheCysArgAspPheG1uLeuG1uLysValCysTrpLeuProLeuAsnThr 680
Db 2070 CTGGAAACTTCTGCAAGACTTGAAGTGAAGAGTGTGTTACTTACCGCTCAACACC 2129
Qy 681 PheLeuLeuArgProLeuH1sArgLeuMetH1sTrYLysGInValLeuG1uArgLeuCys 700
Db 2130 TTCTCTCTCGGGCACTGACCGGCTCATGCACTCAACACAGTCTCGAAGGGCTGTG 2189
Qy 701 LysH1sH1sProProSerH1sL1aAspPheArgAspCysArgL1aL1aLeuL1aG1uL1e 720
Db 2190 AAAACACACCCCGGAGCCGACCTTCAAGGACTCGCGGAGCCCTTTGGCAGAGATC 2249
Qy 721 ThrGluMetValG1uLeuH1sG1YThrMetL1eLysMetGluAsnPheG1uLysLeu 740
Db 2250 ACAGAGATGATGACAGCTCCAGGTACGATGATCAAGATGAGAAATTCAGAAAGCTG 2309
Qy 741 H1sG1uLeuLysLysAspLeuL1eG1YL1eAspAsnLeuValProG1YArgG1uPhe 760
Db 2310 CAGGAATCAAGAAAGATTTGATTGGCAATTGAACAATCTGTGGTTCCGGGAAGGAGTTTC 2369
Qy 761 L1eArgLeuG1YSerLeuSerLysLeuSerG1YLysG1YLeuGInG1uArgPhePhe 780
Db 2370 ATCCGTCTGGGAGCTTCAAGAACTCTCGGGGAAGGGGCTCCAGCAGCGCAATGTTCTTC 2429
Qy 781 LeuPheAsnAspValLeuLeuTrpThrSerArgG1YLeuThrAlaSerAsnGInPheLys 800
Db 2430 CTGTTCAAGAGAGTCTGCTATACAGAGCCGGGGGCTAGCGGCTTCAATCAAGTTTAA 2489
Qy 801 ValH1sG1YL1eLeuProLeuTrpG1YMetThrL1eG1uG1uSerG1uAspG1uTrpG1Y 820
Db 2490 GTTCAAGGAGAGTCCCGCTATGAGCAAGATTTGAAGAGAGCAAGAGCAAGTGGGG 2549
Qy 821 ValProH1sCysLeuThrLeuArgG1YGInArgGInSerL1eL1eValL1aL1aSerSer 840
Db 2550 GTTGGCCCACTGGCTGACCTTCGGGGGCGAGCGAGTTCATCATCTGCGCGGCAAGTTCT 2609
Qy 841 ArgSerG1uMetG1uLysTrpValG1uAspL1eG1uMetAl1eAspLeuL1aG1uLys 860
Db 2610 CGGTCCGAGATGAGAAAGTGGGTGAGAGCAATCCAGATGACCTTACCTGGCGAGAAAG 2669

Qy 861 SerSerSerProAlaProGluPheLeuAlaSerSerProProAlaAsnLysSerProAsp
|||
Db 2670 AGCAGCAGCCCCCGCCCTGAGTTCCTGGCCAGACGCCCCCTTGACAAACAAGTCCCTTAT 2729
Qy 881 GluAlaThrAlaAlaAspGlnGluSerGluAspAspLeuSerAlaSerArgThrSerLeu 900
|||
Db 2730 GAACGACCGCGGCTGACCAAGAGTCAAGAGTACCTGAGGCTCCGCAACATCGCTG 2789
Qy 901 GluArgGlnAlaProAlaArgGluValAsnThrMetValHisValCysThrPheArgAsnThr 920
|||
Db 2790 GAGGCCAGGCCCGCCAGCCGCGCAACAAATGGTGCAGTGTGTGGCACCGCAACCC 2849
Qy 921 SerValSerMetValAspPheSerIleAlaValGluAsnGlnLeuSerGluValAsnLeu 940
|||
Db 2850 AGCGTCTTCATGCGTGGACTTCAGCATCCGAGTGGAAATCAGTTGTCTGGAAACCTGCTG 2909
Qy 941 ArgLysPheLysAsnSerAsnGlyTrpGlnLysLeuTrpValPheThrAsnPheCys 960
|||
Db 2910 AGGAAATTCAAAAACAGCAACGGGTGGCAGAAAGCTGTGGTGTTCACAAACTTCTGC 2969
Qy 961 LeuPhePheTrpLysSerHisGlnAspAsnHisProLeuAlaSerLeuProLeuLeuGly 980
|||
Db 2970 CTGTTCTTCTCAATACACACAGGACAAATCATCCCTTGGCAGCTGCTCTGCTCGGC 3029
Qy 981 TyrSerLeuThrIleProSerGluSerGluAsnIleGlnLysAspTrpValPheLysLeu 1000
|||
Db 3030 TACTCGCTCACCAATCCCTCTGAGTCCAGAAACATCCGAAAGACTACGTGTTCAAGCTG 3089
Qy 1001 HisPheLysSerHisValTrpLysPheArgLysGluSerGluTrpThrPheGluArgTrp 1020
|||
Db 3090 CACTTCAAGTCCCAAGCTTACTACTTCAAGGGCGGAAACGAGTACAGTTGAAAGGTGG 3149
Qy 1021 MetGluValIleArgSerAlaThrSerSerAlaSerArgProHisValLeuSerHisLys 1040
|||
Db 3150 ATGAAAGATCCCGAGTGCACCAAGCTCTGCTGCGAACCCAGTGTGATGTCACAAA 3209
Qy 1041 GluSerLeuValTyr 1045
|||
Db 3210 GAGTCTCTTGTGTAT 3224
RESULT 3
US-10-450-763-564
; Sequence 564, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 564
; LENGTH: 3187
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIMILAR
; LOCATION: (49)..(1371)
; OTHER INFORMATION: 92% homologous to Homo sapiens CDEP, accession number
; OTHER INFORMATION: AB008430, Smith-Waterman Score=2090.
US-10-450-763-564
Alignment Scores:
Pred. No.: 0 Length: 3187
Score: 5431.00 Matches: 1042
Percent Similarity: 99.6% Conservative: 0

Best Local Similarity: 99.6% Mismatches: 3
Query Match: 99.4% Indels: 1
DB: 10 Gaps: 0
US-09-555-342b-2 (1-1045) x US-10-450-763-564 (1-3187)
Qy 1 MetGluValIleGlnGlnArgProThrProGluSerArgLeuGluValAlaProGluAsnSer 20
|||
Db 49 ATGGAGAAATAGACAGAGGCCGACCCAGAGATCAAGACTGGGGGCCCCGAAATTCG 108
Qy 21 GlyIleSerThrLeuGluArgGlyGlnLysProProProThrProSerGluLysLeuVal 40
|||
Db 109 GGAATCAGTACTTGGAAACGTGACAGAAAGCCGCCCAACACTTCAGGAAACTCGTG 168
Qy 41 SerIleValIleGlnMetLeuAspAspThrGlnGlnLysPheGlnValProGlnArgAla 60
|||
Db 169 TCCATCAAAATCCAGATGCTGTGATGACCCAGAGAGGATTTGAAATTCACAAAGACT 228
Qy 61 ProGluLysValLeuLeuAspAlaValCysAsnHisLeuAsnLeuValGluGluAspTyr 80
|||
Db 229 CTTGGAGAGTGTGCTGCTGATGATGCAAGTTTGCAACCACTCAAGCTCTGTGAGAGTGACTAT 288
Qy 81 PheGluLysGluPheProAspHisLysLysIleThrValTrpLeuAspLeuLysPro 100
|||
Db 289 TTTGGCTCGAGTTTCTTGATCAAAAAGATCAAGGTGTGGTGGATCTCTAAACCC 348
Qy 101 IleValLysGlnIleArgArgProLysHisValValValLysPheValValLysPhePhe 120
|||
Db 349 ATTGTGAACAGATTGAAGGCCAAAGCACTTTGTGTAAGTTGTGTGAAATTTCTTT 408
Qy 121 ProProAspHisThrGlnLeuGlnGluGluLeuThrArgTrpLeuPheAlaLeuGlnVal 140
|||
Db 409 CCGCTTACCAACAACAATCTCAAGAAAGACTCAAGAGTCTCTTGGCTGCGCTCAGGTG 468
Qy 141 LysGlnAspLeuAlaGlnArgLysLeuThrCysAsnAspThrSerAlaAlaLeuLeuIle 160
|||
Db 469 AAGCAGAGCTTGGCTCAAGGAGGTTGACGTTATATGACACCAAGCGACTCTTGATT 528
Qy 161 SerHisIleValGlnSerGluIleGluLysPheAspGluAlaLeuAspArgGluHisLeu 180
|||
Db 529 TCACACATTGTCATCTGACATTGGGATTTTATGTAAGCTTGGACAGAGACACTTA 588
Qy 181 AlaLysAsnLysTrpIleProGlnGlnAspAlaLeuGluAspLysIleValGluPheHis 200
|||
Db 589 GCAAAATATATATCATCTTCAACAGACCACTTAAGACAAAATCTGAAATTTTAC 648
Qy 201 HisAsnHisIleGlnIleThrProAlaGluSerAspPheGlnLeuLeuIleAlaArg 220
|||
Db 649 CATAACCACTTGAACAACAACGACGAGAATCAATTTCCAGTCTTAGAATTGCCGT 708
Qy 221 ArgLeuGluMetTrpGlyIleArgLeuHisProAlaLysAspArgGluGlyThrLysIle 240
|||
Db 709 CGGCTAAGATGATGGAATCCGGTGTGACCCGCGCAAGGACAGAGAAAGCAGCAAGATC 768
Qy 241 AsnLeuAlaValAlaAsnThrGlyIleLeuValPheGlnGluPheThrLysIleAsnAla 260
|||
Db 769 AATCTGGCGTTGGCAACGCGGAATTCATGTTTCAAGGTTTCACTAAGATCAATGCC 828
Qy 261 PheAsnTrpAlaLysValArgLysLeuSerPheLysArgLysArgPheLeuIleLysLeu 280
|||
Db 829 TTCAACTGGGCGCAAGGTGGGAACTGAGCTTCAAGAGAAAGCGCTTCTCATCAAGTCC 888
Qy 281 ArgProAspAlaAsnSerAlaTrpGlnAspThrLeuGluPheLeuMetAlaSerArgAsp 300
|||
Db 889 CGGCAAGATGCAATATGCTATCAGATACCTTGAATTCCTGAATGCGCAGTGGAGAT 948
Qy 301 PheCysLysSerPheTrpLysIleCysValGluHisHisAlaPhePheArgLeuPheGlu 320
|||
Db 949 TTCGCAAGTCTCTGGAATCTGTGTGAAATCATGCTCTTTAGACTTTTGAA 1008
Qy 321 GluProLysProLysProLysProValLeuPheSerArgGluSerSerPheArgPheSer 340
|||
Db 1009 GAGCCCAAAACCAAGCCCAAGCCGCTCTTATGCGGGGTGATCATTTCCGTTGAGT 1068

QY 341 G1YArGThrGlnuysGlnValleuAspTyrValIysGluGlyIhIeSLySLyValGln 360
DB 1069 GGTGGAGCTCAAGACAGGTTCTGCACTATTCTAAAGAGAGACATTAAGAGGTCCAG 1128
QY 361 PheGluArgIyShIeSerIyIhIeSerIleArgSerleuAlaSerGlnProThrGlu 380
DB 1129 TTGAAAGAAAGCACACAGATTCAATTCTATCCGGAGCCTTGCTTACAGCCTACAGAA 1188
QY 381 leuAsnSerGluValleuGluGlnSerGlnGlnSerThrSerleuThrPheGluGlu-G1 400
DB 1189 CTGAATTCTGAAAGCTGAGACAGTCTTCAGACAGACCCAGCCTTACATTGGAGAGG 1248
QY 400 YAlaGluSerProGlyGlnSerCyArArgGlyLyseGluProIyValSerAlaG1 420
DB 1249 TCCGGAATTTCCAGGGGCCAGAGCTCCGGCCAGAAAGAACCAAGAGTTCCGGCCG 1308
QY 420 YGluProGlySerThIaProSerProAlaProArgArgSerProAlaGlyAsnLyseGlnAl 440
DB 1309 GGAGCCGGGGTCCGACCCGAGCCTCGCGCCGAGAGAACCCCGGGGTAAACAAGCAGGC 1368
QY 440 aaPrgLyAlaAlaSerAlaProThrGluGluGluGluValValIyShAspArgThrG1 460
DB 1369 GGAACGAGCCGCTCGAGGCCAGAGAGAAAGAGAGAGGTGCTTAAGATAGAGCCCA 1428
QY 460 ngInSerIyProGlnProProGlnProSerThrGlySerleuThrGlySerProhIe 480
DB 1429 GCAGAGTAAACTCAGCCCTCGACGCCAGACACAGGCTCCCTGATCGAGTCTTCACT 1488
QY 480 uSerGluLeuSerValIeAsnSerGlnGlyIyValAlaProAlaAsnValThrLeuSerPr 500
DB 1489 TTCCGAGCTGTGTGAACCTCGCAGGGGAGTGGCCCTGCCAACCTGACTTGTCTCC 1548
QY 500 aaenLeuSerProAspThrIySGlnAlaSerProleuIleSerProleuAsnAspG1 520
DB 1549 CAACCTAGCCCTCGACACCAAGCAGCCTCTCTCTTAATCAGCCGCTGCTGAATGACA 1608
QY 520 naIaCyProArGThrAspAspGluAspGluGlyArGArgLyAspPheProThrAspLy 540
DB 1609 GGCTGCCCCCGAGCGAGCATGAGATGAGGGCCGAGAGAGATCCCAACTGATTA 1668
QY 540 baIaTyrPheIleAlaIySGluValSerThrThrGluArgThrTyrLeuIyAspLeuG1 560
DB 1669 AGCCTACTTCAATAGCTAAAGAAAGTGTCTACACCGAGCAACATATCTGAAGGATCTCGA 1728
QY 560 uValIleThrSerThrPheGlnSerThrValSerIyGluAspAlaMetProGluAla 580
DB 1729 AGTTATCACTTCGTGTTTCAAGACACAGTGAAGAAAGAGACGCCATGCCGAGAGCACT 1788
QY 580 uIySerIeulIePheProAsnPheGluProleuIhIeIyShPheIhIeThrAsnPheLeuLy 600
DB 1789 GAAAGCTCATATNTCCGAAATTTTGAACCTTTGCAACAATTTCTAAATTTTCTCAA 1848
QY 600 sGluIleGluGlnArgleuAlaLeuTyrGluGlyArgSerAsnAlaGlnIleArgAspTy 620
DB 1849 GGAATTTGACCAACGACTTCCCTGCTGGAAAGCCGCTCAAAATGCCAAATTCAGAGATT 1908
QY 620 rGlnArgIleGlyAspValMetleuIyAsnIleGlnGlyMetIyShIeIleuAlaIhIe 640
DB 1909 CCAAGAGATGGGATGTCAATGCTGAAGAACATTCAAGAGCATGAAGCACTGGGGGTCTCA 1968
QY 640 sLeuTyrIyShIeSerGluAlaLeuGluAlaLeuGluAsnGlyIleIySerSerArgAr 660
DB 1969 CTTGTGAACACAGCGAGGCTTTGGAGGCCCTCGAAGATGAAATCAAGAGCTCCGGCG 2028
QY 660 gLeuGluAsnPheCyArArgAspPheGluLeuGlnIyValCyArTyrLeuProleuAsnTh 680
DB 2029 GCTGGAAGAACTTGTGAGAGACTTTGAGCTGCAAGAAAGTGTGTAACTTACCGCTCAAC 2088
QY 680 rPheLeuLeuArgProleuIhIeArgleuMetIhIeTyrIyGlnValleuGluArgleuCy 700
DB 2089 CTTCTCTCTGGGCACTGACCGGCTCATGCACTTAACAAGAGGTCTCTGAGCGGCTGTG 2148

QY 700 sLyShIeAspProSerThIeAlaAspPheArgAspCyArArgAlaAlaLeuAlaGluI1 720
DB 2149 CAACACACACCCGCGAGCCAGCCGACTTCAAGGACTGCGAGCGGCTTTGGCAGAGAT 2208
QY 720 eThrGluMetValAlaGlnleuIhIeGlyThrMetIleIyMetGluAsnPheGlnIySle 740
DB 2209 CACGAGATGTGTGACAGCTTCCACGGTACGAGATCAAGATGAGAAATTTTCAAGAGCT 2268
QY 740 uHIsGluLeuIyAspPheulIeGlyIleAspAsnLeuValValProG1YArGluPh 760
DB 2269 GCACGAATCAAGAAAGATTTGATTTGACATTGACAAATCTTGTTGCTCCGGAAAGGAGTT 2328
QY 760 eIleArgleuGlySerleuSerIyAsnSerGlyIyGlyLeuGlnGlnArgMetPhePh 780
DB 2329 CATCCGTCTGGCAGCCTCAGCAAGCTTCGGGGAAAGGGGCTCCAGACGCGATGTTCTT 2388
QY 780 eLeuPheAsnAspValleuLeuTyrThrSerArgGlyLeuThrAlaSerAsnGlnPheLy 800
DB 2389 CCTGTTCAACGAGTCTCTATACAGAGCCGGGGGCTGACGCTCCAAATCAGTTTAA 2448
QY 800 sValhIsGlyGlnLeuProleuTyrGlyMetThrIleGluGluSerGluAspGluTyrG1 820
DB 2449 AGTCCAGGSCAGCTCCGCTCTATGACATGACGATTTGAGAGAGAGAGACAGATGGG 2508
QY 820 YValProhIsCyAsnLeuThrleuArgGlyGlnArgGlnSerIleIeValAlaIaSerSe 840
DB 2509 GGTGCCCACTGCTGACCTCCGGGGCCAGGGGAGTCATCATGCTGGCGCCAGTTTC 2568
QY 840 rArgSerGluMetGluIySTyrValGluAspIleGlnMetAlaIleAspLeuAlaGluLy 860
DB 2569 TCGGTCCGAGTGGAGAAAGTGGTTGAGACATCCAGATGGCCATTGACCTTGGCGAGAA 2628
QY 860 sSerSerSerProAlaProGluPheleuAlaSerSerProProAspAsnLySerProAe 880
DB 2629 GAGCAGAGCCCGCCCTGAGTTCCTGGCCAGCAGCCCTTGACACAAAGTCCCTGA 2888
QY 880 rGluAlaThrAlaAlaAspGlnGluSerGluAspAspLeuSerAlaSerArgThrSerLe 900
DB 2689 TGAAGCACCGGGGCTGACACAGAGTCAAGAGTGACTGAGGCGCTCGGACATCGCT 2748
QY 900 uGluArgGlnAlaProhIeArgIyAsnThrMetValhIsValCyArTyrhIeArgAsnTh 920
DB 2749 GAGGCCCGAGCCCGCCACCGCGCAACCAATGTGCACTGTGCTGGCAGCGCAAC 2808
QY 920 rSerValSerMetValAspPheSerIleAlaValGluAsnGlnLeuSerGlyAsnLeu 940
DB 2809 CAGCGTCTCATGTGTGACTTCAGCATTCGACGTGGAAATCACTGTCTGGAAACCTGCT 2868
QY 940 uArgLyAspLeuIyAsnSerAsnGlyTyrGlnIySleuTyrValValPheThrAsnPheCy 960
DB 2869 GAGGAATTCAAAACACGACAGGGTGGCAGAAAGTGTGGGTGTTCACAAACTTCTG 2928
QY 960 sLeuPhePheTyrIySerThIeGlnAspAsnIhIeProleuAlaSerleuProleuLeuG1 980
DB 2929 CCTGTTCTTCAAAATCAACACGAGCAATATCCCTTGGCCAGGCTGTGCTCGG 2988
QY 980 YTyrSerleuThrIleProSerGluSerGluAsnIleGlnIyAspTyrValPheIySle 1000
DB 2989 CTATCGCTCACCATCCCTCTGAGTCCGAAACATCCGAAAGACTACGTGTTCAGACT 3048
QY 1000 uHIsPheLySerhIsValTyrTyrPheArgAlaGluSerGluTyrThrPheGluArgTr 1020
DB 3049 GCACTTCAAGTCCCAAGTCTTACAGGGGGGAAAGGAGTATACAGTTCCAAAGGTG 3108
QY 1020 pMetGluValIleArgSerAlaThrSerSerAlaSerArgProhIeValleuSerhIeSly 1040
DB 3109 GATGGAAGATGATCCGAGATGCAACGAGCTGCTCGACCCACAGCTGTGTGAGCACAA 3168
QY 1040 sGluSerleuValTyr 1045
DB 3169 AGAGTCTTGTGTAT 3184

RESULT 4

```
US-10-172-118-1536
; Sequence 1536, Application US/10172118
; Publication No. US20030224374A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linley, Peter
; APPLICANT: Mao, Mao
; APPLICANT: Roberte, Chris
; APPLICANT: Van 't Veer, Laura
; APPLICANT: Van de Vijver, Marc
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-175-999
; CURRENT APPLICATION NUMBER: US/10/172.118
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/380,770
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 1536
; LENGTH: 3997
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: NM_014808
; DATABASE ENTRY DATE: 2001-06-18
US-10-172-118-1536

Alignment Scores:
Pred. No.: 2,966-273 Length: 3997
Score: 2941.50 Matches: 580
Percent Similarity: 70.0% Conservation: 162
Best Local Similarity: 54.7% Mismatches: 261
Query Match: 7 Gaps: 14

US-09-555-342b-2 (1-1045) x US-10-172-118-1536 (1-3997)

QY 1 MercGlyLuiLlegLuiGlnArg-----ProThrProGlySerArgLueGluVal 16
Db 118 ATGGGGAGATAGAGAAATACAGAGTCTGCAGATCGAGGATCGCTTGCGTCC 177
QY 17 ProGluAenSerGlyLieserTherLueGluArgGlyGlnysProProThrProSer 36
Db 178 CAGACCCCTGGGAGATTAGACCTTGAGCTGGGAGACTCTTGCCAGATGCA 237
QY 37 GlyLeuValSerLysLysLysLysLysLysLysLysLysLysLysLysLysLys 56
Db 238 GAGAAACACCTGACCTCAGAGTAAAGCTGCTGGCAACACCATGAAATATTGACATT 297
QY 57 ProGluArgAlaProGlyLysValLeuLeuAspAlaValCysAsnHisLeuAsnLeuVal 76
Db 298 GAGCTTAAATGCGATGGCCAGGTATTACTGCACAGAGTGGAAGCGTTTAACTGGTA 357
QY 77 GluGlyAserProLysPheGlyLysLysLysLysLysLysLysLysLysLysLysLys 96
Db 358 GAATGTACTACTCTGGGATGAGATTTCATAAATCTAGCTCTGATTTGGCTTGA 417
QY 97 LeuLeuValSerProLysLysLysLysLysLysLysLysLysLysLysLysLysLys 116
Db 418 CCTATGAACCCCATCTAGGCAATACGAAAGCCAAAGATGTGGCTTCGCGTGA 477
QY 117 ValLysPhePheProProAspHisThrGlnLueGlnGlnGlnGlnGlnGlnGlnGln 136
Db 478 GTAAAAATTTTTCACCTGATCTGCTCAGCTACAGAAAGAAATATACAAAGTATTGTT 537
QY 137 AlaLeuGlnValLysGlnAserProLysLysLysLysLysLysLysLysLysLysLys 156
Db 538 GCTTGGCACTTAAGAGACCTCTGGAAGAGCTTTGACTGTGCTGACACACAGCG 597
QY 157 AlaLeuLeuLysSerHisLysLysLysLysLysLysLysLysLysLysLysLysLys 176
Db 598 GCCCTTCTACGCTCCATCTCTGCTGAGTGAATTAAGAGATTACGATGAACCTGGAC 657
```

```
QY 177 ArgGluHisLeuAlaLysAsnLysTyrLysProGlnGlnAspAlaLeuGluAspLysLys 196
Db 658 CGAAGGACCTTCAAGAGATATTGGCTGGCCAGACGACTGCTTGAAGAGATA 717
QY 197 ValGluPheHisHisAsnHisLysGlnThrProAlaGlnSerAspPheGlnLeuLeu 216
Db 718 CTAGATTTCCATCAGAGACAGCTGGCCAGACACTGCTGAGTGGATTTCAGGTGCTC 777
QY 217 GluLysAlaArgArgLysLysLysLysLysLysLysLysLysLysLysLysLysLys 236
Db 778 GAAATTCCTCGAAAGTTGGAATTAAGGCTATGATTTCAACAGGCTTCTGACAGGGAA 837
QY 237 GlyThrLysLysLeuLeuAlaValAlaAsnThrGlyLysLeuValPheGlnGlyPheThr 256
Db 838 GAAACCAAGATTCAACTGGAGTTTCCACATGGGTACTCGTGTCCAGAGGACCAAC 897
QY 257 LysLysAsnAlaPheAsnThrProLysValArgLysLysSerPheLysArgLysPhe 276
Db 898 AAATCAACACTTTCATCTGCTGCTCAAGTCCGTAACCTTAAGCTTCAAGAGAAAGATT 957
QY 277 LeuLysLeuArgProAspAlaAsnSerAlaTyrGlnAspThrLeuGluPheLeuMet 296
Db 958 CTATCAAACTTCATCCAGAGGTTCAATGACCTTACAGACACATTAGATTATTTGTG 1017
QY 297 AlaSerArgAspPheCysLysSerPheThrLysLysLysValGlnHisAlaPhePhe 316
Db 1018 GGTATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1077
QY 317 ArgLeuPheGlnGluProLysProLysProLysProValLeuPheSerArgLysSerSer 336
Db 1078 AGATTTTTGGACCAACTTAAGCCAAAGAAAGACCGCTTCTTCAAGCGGGGCTCTCC 1137
QY 337 PheArgPheSerGlyArgThrGlnLysGlnValLeuAspTyrValLysGlnGlyHis 356
Db 1138 TTCAGATACAGTGGAAAGAACTCAGAAACAACTAGTATGATTATTTCAAGACAGTGAATG 1197
QY 357 LysLysValGlnPheGlnArgLysLysLysLysLysLysLysLysLysLysLysLys 374
Db 1198 AAGAAATTCATATGAAAGAGGACAGACAGACCAACGCTCGTTCAGCTTGA 1257
QY 375 AlaSerGlnProThrGlnLeuAenSerGlnValLueGlnGlnSerGlnSerThrSer 394
Db 1258 GCACACTTACCAAAACAG-----AGC 1278
QY 395 LeuThrPheGlyGlnGlyAlaGlnSerProGlyGlnSerCysArgArg----- 411
Db 1279 ATCTCATTTCCCGAGGATGAGGACTCTGCTCCCATCTTCCAGCGAATGCTTTTAC 1338
QY 412 -----GlyLysGlnProLysValSerAlaGlyLysProGly----- 423
Db 1339 TCGCTCTCTCTCCACTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1398
QY 424 -----SerHisProSerProAlaProArgArgSerProAlaGlnLysGlnAlaAsp 441
Db 1399 TCCCTCACAGATCCCGAGGTTTCTTCAAGATCAGCTGACAGAGAGCGGAGGGA 1458
QY 442 GlyAlaAlaSerAlaProThrGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 461
Db 1459 GCAATGCTGAGAGGCGCCGAC-----ACACCATTCG 1488
QY 462 SerLysProGlnProPro-----GlnProSerThrGlySerLeuThrGlySerPro 478
Db 1489 GCCAGGCTCTCGGGGCCCCCGGACCTCCAGCTGCTGCTGCTTCCAGAGAGATCT 1548
QY 479 HisLeuSerGlnLeuSerValAsnSerGlnGlyGlyValAlaProAlaAsnValThrLeu 498
Db 1549 CAGCTTCTCTCCCTCCAGCGGAGAGCGCCCTGAGTGAAGCTTGA-----TTT 1599
QY 499 SerProAsnLeuSerProAspThrLysGlnAlaSerProLeuLysLysSerProLeuAsn 518
Db 1600 CAGGTGCTTTGGGCGCAGCTGAAACAGGGCTCATCTCCACTCTCGAGACCTGCTCAGT 1659
```

Qy	519	AAAGGAlaAeYsrProArghTArBarAgUaArBgInUGUaXgUaGlyArBarProth	538
Db	1660	GAATGGTGGGCGAGCCGGAGATGAGCTGGAG--GAGCCGAGACAGAGGCTGGCTCA	1716
Qy	539	AspUyAlaYrPheIleAlaUyGluValSerThrThiGluArgThrTyTLeuUyAsp	558
Db	1717	GAGAGGCGCTACTTCATAGTCAGAAAGAGATTTCTCGTCAAGAAAGAACTACTCAAGAT	1776
Qy	559	LeuGluValIleThiSerThrPheGlnSerThrValSerUyGluAspAlaMetProGlu	578
Db	1777	TTGAAAGTTATTACCGGTGGTCTCCGACAGCGAGTGTGAAGAGAGAGCCATGCTCGC	1836
Qy	579	AlaLeuUySerLeuIlePheProAspPheGlnProLeuIleUyPheIleThiAspPhe	598
Db	1837	ACTTAAATGAGCGTGGCTTTCTTCAGATCAATGAGCCATTAATGAGTTCCACAGAGCTTC	1896
Qy	599	LeuUyGlnUleGluGlnArgLeuAlaLeuThrPGLUUAArgSerAspAlaGlnIleArg	618
Db	1897	CTGCGGAGGTGAGACAGAGGCTGGGACATCTGGGAAAGGGCCCTCCAAAGCCAGCAAA	1956
Qy	619	--AspUyGlnArgIleGlyAspValMetLeuUyAspIleGlnUyMetUyIleU	637
Db	1957	GCGAGTCATCAAGAAATCGGGAGCATCTGCTCAGAAATGCGCAAGTAAAGAGTTT	2016
Qy	638	AlaAlaIleValLeuThrUyIleAspGlnAlaLeuGlnAlaLeuGlnUyIleUySer	657
Db	2017	ACCAAGTACTTCCAAAGACATGACGAGGTCTTAAAGAACTGAAAGGCTCAAAAGCC	2076
Qy	658	SerArgArgLeuGlnUyAspPheCyArAspPheGlnUyGlnUyValCytUyLeuPro	677
Db	2077	TGTAAAGATGTGAGGACAGTGTACAGAGATTTGAAGTGTCAAGAGTCTGTACTTGGCT	2136
Qy	678	LeuAspThrPheLeuUyArgProLeuUyIleArgUyMetUyIleUyGlnUyAlaUyGln	697
Db	2137	CTCAACACGTTCTCGCTGAAGCCCATCCAGCGGCTGTGCACTACGCGCTGTCTGGCGC	2196
Qy	698	ArgLeuUyCysUyIleIleIleProProSerIleValAspPheArgUyAspArgAlaIleU	717
Db	2197	CGGCTTAGCGGACATTACAGCCCCGGGACCATGATGATGAGTGGCTGATGACGCCCTG	2256
Qy	718	AlaGlnUleThrGlnUyMetValAlaGlnUyIleUyIleMetIleUyMetGlnUyAspPhe	737
Db	2257	AAAGGCTTCAAGAGGTGACACCAACATCAAGACATTTCTATCCGCTGGAGAACTG	2316
Qy	738	GlnUyLeuUyIleGlnUyUyUyUyAspLeuUyIleGlyIleAspAsnUyValValProGly	757
Db	2317	CAGAGCTTAACGAGGCTGACGAGGGAACGTGATGAGGACATGAGAACTCATTTCTCTCGC	2376
Qy	758	ArgGlnUyPheIleArgLeuUySerLeuUySerUyUyUyGlyUyGlyUyGlnUyGlnUyArg	777
Db	2377	AGGAGATTACTCCGTGAGGCTGCTCTTCAAGCTCAACMAAGAGGCTGTGACAGAGG	2436
Qy	778	MetPhePheUyPheAspValIleUyUyUyThrSerArgGlyUyLeuThrAlaSerAsp	797
Db	2437	ATGTTTTTTTGTCTTCAAGTATATGTTGCTGTACACAGAAAGAGATTGACAGGACGAC	2496
Qy	798	GlnPheUyValIleGlyGlnUyLeuProLeuUyUyGlyUyMetThrIleGlnUySerUyAsp	817
Db	2497	CACCTCCGATCCGGGGGCTCTCTCCCTCCAGAGCATCTGGTGAAGAAAGTATAAC	2556
Qy	818	GlnUyTrpGlyValProIleCysUyLeuThrLeuUyArgGlyGlnArgGlnSerIleUyAla	837
Db	2557	GAGTGGTCTGTTCACACTGTTTCAACCATTCACGCGGCTCAAAAAAATCTGTGTGGCA	2616
Qy	838	AlaSerSerArgSerGlnUyMetGlnUyUyTrpValGlnAspIleGlnMetAlaIleAspU	857
Db	2617	GCGACGACTGGCTGGAGAAAGAGATGATGCTGACCTGAATCTCCGCGATCCAGACA	2676
Qy	858	AlaGlnUySerSerSerProAlaProGlnUyPhe-----LeuUyAspSerPro	873
Db	2677	GCGAAGAGTGGGGTGACACGGCCCTGACATCCAGCGCGGACATGTGTGACATCTGTCC	2736
Qy	874	ProAspAsnUySerProAspGlnUyAlaThrAlaIleAspGlnUySerGlnUyAspLeu	893

Db	2737	CCC-----AGATCCCCCAAGAG---GTATCTCTGGAGCAGAGACTCAGAAAGATGATCTT	2787
Qy	894	SeRlaserAqrThrSerLeuGluArgGlnAlaProHisArgGlyAsnThrMetValHis	913
Db	2788	CGGGGTGTCGCGAGCTCCCTCGAGGGGAGTGGCAGCACCGGGGCCAAACACCAATGCAC	2847
Qy	914	ValCyStrpHisArgAsnThrSerValSerMetValAspPheSerIleAlaValGluAsn	933
Db	2848	GTGTGCTGTGTATCCGGAACACCAGCGGTGTCCAGGGCAGACCCACATGCAGCTGTGGAGAAC	2907
Qy	934	GlnLeuSerGlyAsnLeuLeuArgLysPheLysAsnSerAsnGlyTyrPglLysLeuTyr	953
Db	2908	CAGCTTCAGAGATATCTGCTAAGAAAGTCTCAAADACAGTCAGCTGGCCGAAAGCTCTGG	2967
Qy	954	ValValPheThrAspPheCysLeuPhePheTyrLysSerHisGlnAspAsnProLeu	973
Db	2968	GTCTGCTTTACCAATCTCTGTGTCTTCTTCAAAATCATATCAGATGATACCTACCAGT	3027
Qy	974	AlaSerLeuProLeuLeuGlyTyrSerLeuThrIleProSerGluSerGluAsnIleGln	993
Db	3028	GCCAGCGTCCCGGTGCTGGGGCTTACAGGCTAGCATCCCAAGGAGGCGGATGGCATAC	3087
Qy	994	LysAspTyrValPheLysLeuHisPheLysSerHisValTyrTyrPheArgAlaGluSer	1013
Db	3088	AAAGACTATGTTTCAAGCTCCAGTTCAAATCCACACGCTACTCTTCCGGGGCTGAGAGC	3147
Qy	1014	GluTyrThrPheGluArgTyrMetGluValIleArgSerAlaThrSerSerAlaSerArg	1033
Db	3148	AAGTTCACATTGAAAGGTGATGAGGTATCCAGGGGGCCACACAGCTCACCGCGGAGG	3207

```

RESULT 5
US-10-342-887-1536
; Sequence 1536, Application US/10342887
; Publication No. US20040058340A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yundong
; APPLICANT: Linsley, Peter S.
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Christopher J.
; APPLICANT: Van 't Veer, Laura Johanna
; APPLICANT: Van de Vijver, Marc J.
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-188-999
; CURRENT APPLICATION NUMBER: US/10/342,887
; CURRENT FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: 60/238,918
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/380,710
; PRIOR FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: 10/112,118
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 1536
; LENGTH: 3997
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-342-887-1536

Alignment Scores:
Pred. No.:      2,966-273          length:    3997
Score:         2941.50             Matches:     580
Percent Similarity:   70.0%       Conservative:  162
Best Local Similarity: 54.7%       Mismatches:   261
Query Match:        53.8%         Indels:       57
DB:                8              Gaps:         14

US-09-555-342B-2 (1-1045) x US-10-342-887-1536 (1-3997)
QY      1 Metcylgtctgctgcagatgg-----PrcthrProclgyserarglengciayala 16
|||||

```

Db 118 ATGGGGGAGATGAGAAACATACAGAGCTCGTGCAGATCGAGGGAGTCGGTGGGTC 177
Qy 17 ProGUbenSerGlyIleSerThrLeuGluArgGlyGlnIlyProProProThrProSer 36
Db 178 CAGACCCCTGTGGAGATTGACACCTTGGAGCTGGGCGAGCTCTTGGCCAGATGCA 237
Qy 37 GilyLysLeuValSerIleValIleGluMetLeuAspThrGlnGluAlaPheGluVal 56
Db 238 GAGAGGACCTGCACCTCAGAGTAAAGCTGTCGACACACCAAGATTTTGACATT 297
Qy 57 ProGUArgAlaProGlyLysValLeuLeuAspAlaValCysAsnHisLeuAsnLeuVal 76
Db 298 GAGCTAAATGCGATGGCCAGATATTACTGACACAGTGGAGAGCTTTAAACCTGTA 357
Qy 77 GGUgluAspGlyThrPheGlyLeuGluPheProAspHisIlyValIleThrValTrrLeuAsp 96
Db 358 GAATGTGACTACTTCGGGATGAGATTGCAAAATACAGTCTGATGATTTGGCTTGA 417
Qy 97 LeuLeuLysProIleValIleGlnIleArgArgProLysHisValValIleValPheVal 116
Db 418 CTAATGAAACCAATATTAGGCAATACAGAGCCAAAGATGGAGTCTTCGGCTAGCT 477
Qy 117 ValLysPhePheProProAspHisIleThrGlnLeuGlnGluLeuThrArgTrrLeuPhe 136
Db 478 GTAATAATTTTCCACCTGATCCTGGTCAAGTCAAGAAATATACAAAGATACCTGTTT 537
Qy 137 AlaLeuGlnValIlyGlnAspLeuAlaGlnIlyArgLeuThrCysAsnAspThrSerAla 156
Db 538 GCCTTGCAACTTAAAGAGACCTGCTGGAGAGCGCTTTGACCTGGCTGACACACAGCG 597
Qy 157 AlaLeuLeuIleSerHisIleValIleGlnSerGlnIleGlyAspPheAspGlnAlaLeuAsp 176
Db 598 GCCCTTCCACGCTCCCATCTCTCGAGTCGGAAATAGAGATTACAGATGAAACGCTGAC 657
Qy 177 ArgGlnHisLeuAlaLysAsnLysTrrIleProGlnGlnAspAlaLeuGluAspLysIle 196
Db 658 GGAAGGACCTCAAGATGAAAGATATTGGCTGGCCAGACAGACCTGCTTGAGAGATTA 717
Qy 197 ValGluPheHisIleAsnHisIleGlyGlnThrProAlaGluSerAspPheGlnLeuLeu 216
Db 718 CTAAGATTCCATCAGAAAGACAGTGGGCCAGACACCTGAGTGGAGATTTCAGAGTCTC 777
Qy 217 GGUllyLeuArgArgLeuGluMetTrrGlyIleArgLeuHisIleProAlaLysAspArgGlu 236
Db 778 GAATTTGCTCAAGATTTGAAATGATGCGCATCAGATTTCACATGCTTTCGACAGGGA 837
Qy 237 GilyThrLysIleAsnLeuAlaValAlaAsnThrGlyIleLeuValPheGlnGlyPheThr 256
Db 838 GGAACCAAGATTCAACTGGCAGTTCCCAATGGGTGATCTGTTCCAGGGGACCAAC 897
Qy 257 LysIleAsnAlaPheAsnTrrAlaLysValArgLysLeuSerPheLysArgLysArgPhe 276
Db 898 AAAATCAACACTTTCACCTGTCCAAGGTCGTAACCTTCAAGAGGAAAAAGATT 957
Qy 277 LeuIleLysLeuArgProAspAlaAsnSerAlaTrrGlnAspTrrLeuGlnPheLeuMet 296
Db 958 CTAATCAAACTTCATCCAGAGTTCATGACCTTACCCAGACACATTTAGATTGTTGTT 1017
Qy 297 AlaSerArgAspPheCysLysSerPheTrrLysIleCysValGlnHisIleAlaPhePhe 316
Db 1018 GGTGATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1077
Qy 317 ArgLeuPheGlnGluProLysProLysProLysProLysProLysPheSerArgLysSer 336
Db 1078 AGACTTTTGGACCAACCTAACCAAAAGCAAAAGCCGCTTCTTCAAGCCGGGCGCTCC 1137
Qy 337 PheAsnThrPheLeuArgTrrGlnIleGlnValLeuAspTrrValIlyGlnGlyIlyHis 356
Db 1138 TTCAGATCAGTGGAGAACTCAGAAACACTAGATGATTTTTCAGAAAGACGTTGGAATG 1197
Qy 357 LysLysValGlnPheGluArgLysHisIleSerLysIleHisIleSerLysSerLeu 374
Db 1198 AAGAGATTCATATGAAAGAGGACAGCAAGACCAACGACGTCGTTGAGCTCTGACT 1257

Qy 375 AlaSerGlnProThrGlnLeuAsnSerGluValLeuGlnGlnSerGlnIleThrSer 394
Db 1258 GCAGACTTACCAAAACAG-----AGC 1278
Qy 395 LeuThrPheGlyGlnGlyAlaGluSerProGlyGlyGlnSerCysArgArg----- 411
Db 1279 ATTCATTTCCCGAGGAGATTTAGAGATCTCTCCCTCCCATCTTCAGGAAATGCTTTTAC 1338
Qy 412 -----GlyLysGluProLysValSerAlaGlyLysProGly----- 423
Db 1339 TCGCTCTCTCCCTCAGCTGTGTCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1398
Qy 424 -----SerHisAspSerProAlaProArgArgSerProAlaGlyAsnLysGlnAlaAsp 441
Db 1399 TCCCTCAGAGATCCCGAGATTTCCTTACGTCAAGATCAGCTGACAGAGAGCGAGTGA 1458
Qy 442 GilyAlaAlaSerAlaProThrGlnGluGlnGluValIlyLysAspArgThrGlnGln 461
Db 1459 GCAGTGGCTGAGAGCCCGAC-----ACACCATCG 1488
Qy 462 SerLysProGlnProPro-----GlnProSerThrGlySerLeuThrGlySerPro 478
Db 1489 GCCCAGCCCTCGGGCCCCCGACCTCAGGCTGTCAGGCTTTCACAGAAAGATCCT 1548
Qy 479 HisLeuSerGluLeuSerValAsnSerGlnGlyIlyValIleProAlaAsnValThrLeu 498
Db 1549 CAGCTTCTCTCCCTCAGCCGAGAGAGCCCTGAGTCTGAGCCCTGCA-----TTT 1599
Qy 499 SerProAsnLeuSerProAspThrLysGlnAlaSerProLeuIleSerProLeuAsn 518
Db 1600 CAGTGGCTTTGGGCCAGCTGAAACAGGGCTCATCCCACTCCGAGGCCCTGCTCAGT 1659
Qy 519 AspGlnAlaCysProArgThrAspAspGlnAspGlnIlyArgArgLysArgPheProThr 538
Db 1660 GATCTGCGGAGCGGAGCGGAGATGAGCTGCGAG---GAGCCGACACCAAGCGGCTCGCA 1716
Qy 539 AspLysAlaTrrPheIleAlaLysGlnValSerThrThrGlnArgTrrTrrLysLysAsp 558
Db 1717 GACGAGGCTTACTTCAATGATCAAGAAAGATTTCTCGTCAAGAAACCAATACCTTAGAT 1776
Qy 559 LeuGluValIleThrSerTrrPheGlnSerThrValSerLysGlnAspAlaMetProGlu 578
Db 1777 TTAAGATTAATTAACGCTGCTGCTTCCGACGCCATGTTGAAGAGACCATCTCGG 1836
Qy 579 AlaLeuLysSerLeuIlePheProAsnPheGlnProLeuHisIlyPheHisIleAsnPhe 598
Db 1837 ACTGTGATGACGCTGCTCTTCCAAATGATGATCCATCTATGATTCACAGAGGCTTC 1896
Qy 599 LeuLysGlnIleGlnGlnArgLeuAlaLeuTrrGlnGlyArgSerAsnAlaGlnIleArg 618
Db 1897 CTGGCGAGGTTGAGGAGAGGCTGCGACTGCGGAGGCGCTCCAAAGCCACACAA 1956
Qy 619 ---AspTrrGlnArgIleGlyAspValMetLeuLysAsnIleGlnIlyMetLysIleLeu 637
Db 1957 GGCAGTATCAACAAATCGGGGACATCTCTGTCAAGAAATGCCCCAGTTAAAGAGTTT 2016
Qy 638 AlaAlaHisLeuTrrLysHisIleSerGluAlaLeuGlnAlaLeuGlnLysIleLysSer 657
Db 2017 ACCAGTACTTCCAAAGACATGACGAGTCTTAACAGAACTGGAAGAGGCTAACCAAGCGC 2076
Qy 658 SerArgArgLeuGlnLysAsnPheCysArgAspPheGlnGlnLysValCysTrrLeuPro 677
Db 2077 TGTAAAGATTGAGGAGCTGTAACAGAGATTGAGTGGAGAAAGTCTGTACTTGTGCT 2136
Qy 678 LeuAsnThrPheLeuLeuArgProLeuHisIleArgLeuMetHisIlyLysGlnValLeuGln 697
Db 2137 CTCAGACGTTCTCGTGAAGCCATCCAGGGGTGTGATGACATCCGCGCTGCTGCGC 2196
Qy 698 ArgLeuCysLysHisIleProProSerHisIleAspPheArgArgCysArgAlaAlaLeu 717
Db 2197 CGCCTTAGCGGACATTACAGCCCGGAGCACCATGATACGTGCTGATGAGCCCTG 2256


```

QY      718 AlaGluIleThrGluMetValAlaIleuLeuHisgLIyLhmTMeLiIeLyMeGluAsnPh 737
Db      2257 AAAGCATTACAGAGTAGTCCACCACACTACAGCACATTCCTCATCCGGCTGAAGAAGCTG 2316
QY      738 GlnTyLSueNHsgLIuLeuTyLeuYSArPheuiLegLIleAapAmLeuValProGIy 757
Db      2317 CAGAAAGCTAACCGAACTGCAGGGGAACTGTGGGCAATAGAAAGCTCATTTGCTCTGGC 2376
QY      758 ArgGluPheIIeaArgLeuGIySerLeuSerTyLSueSerGIyLYegLyLeuGInGlnArg 777
Db      2377 AGGGAGTTCATCCGTGAGGGGCTGCTTCACAAGCTCACCAAGAAAGGGGCTCACAAGAG 2436
QY      778 MetPhePheLeuPheAsnAPheValLeuLeuTYrThrsARagGlyLeuThrAlaSerAn 797
Db      2437 AAGTTTTTCTGTCTTCAGATATGTGCTGTACACAAACCAAAGAGGTGAGGAGCAAGC 2496
QY      798 GlnPheLYsVAlHsgLIuLeuProLeuTYrGIyMeTthrIIeGLuGIsuGIuAP 817
Db      2497 CACTTCCGGATCCGGGGGCTCTTCCCCTCCAAGSAlTGCTGTGAAGAAAGTGAATCA 2556
QY      818 GIUTrgLIyValProHICyeLeuThrLeuArgGlyGlnArgGInserIIleValAla 837
Db      2557 GAGTGGTCTGTTCACACACTGTTTACCATCTACCGGGCTCAGAAAACAATGTGTGGCA 2616
QY      838 AlasSerSerArgSerGIuMeGIuLySTrpValGluAPrIIeGLImetAlIlleaPhe 857
Db      2617 GCCACACACTCCGGCTGGAAGAAGAAAGTGAATGTGCTGACTCCGCCATCCAAAGCA 2676
QY      858 AlaGIuLySeserSerProAlaProGIuPhe-----LeuAlaSerSerPro 873
Db      2677 GCCAAGAGTGGCGGTGACACGGCCCCCTGACCTGCCAGGCCCGCATGTGTGACTGTCTCC 2736
QY      874 ProAepAmLySeserProApGIuAlaThzAlalaAepGInGIuSerGIuAsPheu 893
Db      2737 CCC-----AGATCCCCCAAGAG--GNATCTCGAGACAGAGTCAAGAAATGATGCT 2787
QY      894 SerAlaSerArgThrSerLeuGIuArgGlnAlaAlaProHISargLyLeuThrMetValHis 913
Db      2788 CGGGGTGTCCGCAAGCTCCCTCGAGGGGCGATGGCCAGCACCGGGCCACACACCAATGCAC 2847
QY      914 ValCYstrPHIArXArAmThrSerAlaSerMetValaAPheserIIleAlaValGIuAsn 933
Db      2848 GTGTCTGTATCCGGAAACCAAGCTGTCTCCAGGGCAGACACAAAGCAGCTGTGAAGAC 2907
QY      934 GlnLeuSerGIyASnLeuLeuArgLyAPheLySaenSerAangITrPGInLyLeuTrp 953
Db      2908 CAGCTTTCAGGATATCTGCTGAAGAAGTTCAAAAAACAATCATGGTGCACAAGACTCTGG 2967
QY      954 ValValPheThraPhePheCysLeuPhePheTYrLySeriHsgLIuAsnHisProLeu 973
Db      2968 GTCCGCTTTACCAACTTCCTGTTTCTTCTTAACAAACTCATCTCAGAGATACCACTG 3027
QY      974 AlaSerLeuProLeuLeuGIyTyrSerLeuThrIIeProSerGIuSerGIuAsnIIeGln 993
Db      3028 GCCAACCTCCCGCTCTCGGCTACACAGCGTGAAGCATCCCAAGGAGGCCGATGGCAATAC 3087
QY      994 LySAepTYrValPheLyLeuHisPheLySeriHISValITYrThrPheARGAlaGIuSer 1013
Db      3088 AAAAGCTATGTTTTCAGACTCCAGTTCAAATCCCACTCATCTTCTCCGGGCTGAAGAC 3147
QY      1014 GluTYrThrPheGluArgTYrMetGIuValIIeaArgSerAlaThrsSerSerAlaSerArg 1033
Db      3148 AAGTACACATTTGAAAGGTGATGATGATCCAGGGGGCCAGACGCTCAGCGGGAGG 3207

```

```

/ CURRENT FILING DATE: 2004-03-17
/ NUMBER OF SEQ ID NOS: 29
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 16
/ LENGTH: 3997
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: KIA00793
/ LOCATION: (1)..(3997)
/ OTHER INFORMATION: GenBank accession No. AB018336
US-10-802-432-16

Alignment Scores:
Pred. No.: 2,966-273 Length: 3997
Score: 2941.50 Matches: 580
Percent Similarity: 70.0% Conservative: 162
Best Local Similarity: 54.7% Mismatches: 261
Query Match: 53.8% Indels: 57
DB: 9 Gaps: 14

US-09-555-342B-2 (1-1045) x US-10-802-432-16 (1-3997)
QY 1 MetGIyGluIleGluGlnArg-----ProThrProGlySerArgLeuGIyAla 16
Db 118 ATGGGGAGATGAAGAAACATACAGAGTCTGCAGACTGCAGGGATGGCGTTGGGAGCC 177
QY 17 ProGluAseSerGIyIleSerThrLeuGluArgGIyGluIleAspProProThrProSer 36
Db 178 CAGACCCCTGGGGAGTTAGACCCCTTGAGCCTGGGCACACTCTCTGGCCAGAAATGCAA 237
QY 37 GIyIyLeuValSerIleValIleGlnMetLeuAspPheThrGlnGluAlaPheGluVal 56
Db 238 GAGAAAGCACTGCACCTCAGAGTAAAGCTGCTGGACACACATGAGAAATATTGTGACATT 297
QY 57 ProGlnArgAlaProGIyIyValIleLeuAspAlaValCysAsnHisIleuAsnLeuVal 76
Db 298 GAGCCTAAATGGATGGAGGCCAGGATATTACTGACACAAAGTGGAAAGCGTTAAACCTGGTA 357
QY 77 GluGIyAspPhePheGlyLeuGlnPheProAspHisIyLeuAlIethrValTrpLeuAsp 96
Db 358 GAAATGTGACTACTTGGGATGAGATTTCAAATATCTCACTCCTACATGAGATTGGCTTGAA 417
QY 97 LeuLeuIyPheProIleValIyGlnIleArgArgProIyAsnIleValIyValIyAspPheVal 116
Db 418 CCTATGAAACCCATCATTTAGCGAAATTAACGAAGGCCAAAGAAATGTGGTTCCTGACCTAGCT 477
QY 117 ValIyPhePheProProAspHisIthrGlnIleGlnGlnIleuThrArgTrpGlyLeuPhe 136
Db 478 GTAAATATTTTCCACTGATCTTGATCAGTACAAAGAAATATACAAAGATATCTTGTTT 537
QY 137 AlaLeuGlnValIyGlnAspLeuAlaGlnIyArgLeuThrCysAsnAspThrSerAla 156
Db 538 GCCTTGCAACTTAAGAAGAACTGCTGGAAAGACGTTTGAACCTGTGCTGACACCAACAGCG 597
QY 157 AlaLeuLeuIleSerHisIleValGlnSerGlnIleGlyAspPheAspGluAlaLeuAsp 176
Db 598 GCCCTTTCACGTCCTCATCTCTGACAGTGGAAATAGAGAGATTACGATGAAACGCTGGAC 657
QY 177 ArgGluHisIleuAlaIyAsnIyTrIlePheGlnGlnIleAspAlaIleuGluAspIyIle 196
Db 658 CGAGAGCAACCTCAAGTAGAACGAGATTGTGCTGGGCAGACAGCACTGCCCTTGAAAGATA 717
QY 197 ValGluPheHisIleAsnHisIleGlnIleThrProAlaGlnIleSerAspPheGlnIleu 216
Db 718 CTAGAAATTCATCAGAAAGCACTGGGACCAAGACACTGCTGATGGATTTCCAGAGTGCTC 777
QY 217 GluIleAlaArgArgLeuGlnMetTrpGlyIleArgGlnHisProAlaIyAspArgGlu 236
Db 778 GAAATTTGCTCGAAAGTTGAAATGTACGGCATCAGATTTCACATGGCTTCTGACAGGAA 837
QY 237 GIyThrIyValIleAsnLeuAlaValAlaAsnThrGIyIleLeuValPheGlnGlyPheThr 256

```

Db 838 GGAACCAAGATTCAACTGGACGTTCCACATGGGTGACTGCTGTTCCAGGCGACACC 897
Qy 257 LysIleAsnAlaPheAsnTrpAlaValArgLysLeuSerPheLysArgLysArgPhe 276
Db 898 AAAATCAACACTTCACTGCTCAAGGTCGTAACCTTAAGCTTAAGAGGAAAAATT 957
Qy 277 LeuIleLysLeuArgProAspAlaAsnSerAlaTrpGlnAspThrLeuGlnPheLeuMet 296
Db 958 CTTATCAAACTTCATCCAGAGGTTCAAGACCTTACAGGACACTTGAATTTTGTG 1017
Qy 297 AlaSerArgAspPheCysLysSerPheTrpLysIleCysValGlnHisIleAlaPhePhe 316
Db 1018 GGTGTAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1077
Qy 317 ArgLeuPheGlnGluProLysProLysProLysProValLeuPheSerArgLysSer 336
Db 1078 AGACTTTGGACCACTTAAGCCAAAGCAAAAGCCGCTTTTACGGGGGCTCTCC 1137
Qy 337 PheArgPheSerGlyArgThrGlnLysGlnValLeuAspTrpValLysGlnLysLys 356
Db 1138 TTCAGATACAGTGGAGAACTCAGAAACACTAGTATGTTTCAAGACAGTGGATG 1197
Qy 357 LysLysValGlnPheGlnArgLysHisSerLysIleHisSerLysArgSerLeu 374
Db 1198 AAGAGAAATTCATATGAAAGAAAGACAGCAAGACCAAGCTCCGCTCGAGCTGACT 1257
Qy 375 AlaSerGlnProThrGlnLeuAsnSerGlnValLeuGlnInsSerGlnSerThrSer 394
Db 1258 GCAAGACTACCAAAACAG-----AGC 1278
Qy 395 LeuThrPheGlnGlyGlnAlaGlnSerProGlnGlyGlnSerCysArgArg----- 411
Db 1279 ATCTCATTTCCCGAGGAGATTGAGAGCTCTGCTCCCATCTTCAAGCAATCCCTTTAC 1338
Qy 412 -----GlyLysGlnProLysValSerAlaGlyGlnProGly----- 423
Db 1339 TCGCTCTCTCCCTCACTCTGCTCCCTGCTGCGCTGAGATTAAAGACAGCAGCAGC 1398
Qy 424 -----SerHisProSerProAlaProArgArgSerProAlaGlyAsnLysGlnAlaAsp 441
Db 1399 TCCTTCACAGATCCCGACGTTTCTCACTCAAGAGTCCAGCTGACAGAGAGCGCGTGA 1458
Qy 442 GlnAlaAlaSerAlaProThrGlnGlnGlnGlnValLysAspArgThrGlnGln 461
Db 1459 GCACTGGCTGAGGCGCCGAC-----ACACCAACG 1488
Qy 462 SerLysProGlnProPro-----GlnProSerThrGlnLysSerLeuThrGlySerPro 478
Db 1489 GCCCAGCCCTCGGGGCCCGCCGCACTCCAGCTGGTCCAGGCTTTTCCAGAAAGATGCT 1548
Qy 479 HisLeuSerGlnLysSerValAsnSerGlnGlnValAlaProAlaAsnValThrLeu 498
Db 1549 CAGCTTCTCTCCCTCAGCGGAGAGCCCTGAGTCTGAGCCCTGCA-----TTT 1599
Qy 499 SerProAsnLeuSerProAspTrpThrLysGlnLysSerProLeuLysSerProLeuAsn 518
Db 1600 CAGGTGCTTTGGGCGCCAGCTGAAACAGGGCTCATCCCACTCTCGAGGCCCTGCTCAGT 1659
Qy 519 AspGlnAlaCysProArgTrpHisAspAspGlnAspGlnLysArgLysArgPheProThr 538
Db 1660 GATCTGCGCGGAGCGGAGTGAATGAGAG-----GAGCCCAACACAAACGCGTGCCTCA 1716
Qy 539 AspLysAlaTrpPheIleAlaLysGlnValSerThrThrGlnLysGlnLysLysAsp 558
Db 1717 GACCAAGGCTACTCATATGTCAAAGAAATTTCTCGCTACAGAACAAACATACCTTAAGAT 1776
Qy 559 LeuGlnValIleThrSerTrpPheGlnSerThrValSerLysGlnAspAlaMetProGln 578
Db 1777 TTAAAGATTATTACCGTGTGTTCCGCGAGCGAGTGGTGAAGAGAGAGCCATGCTCGG 1836
Qy 579 AlaLeuLysSerLeuIlePheProAsnPheGlnProLeuHisLysPheHisThrAsnPhe 598
Db 1837 ACTGTATGACGTGCTCTCTCCAAACATCGATCCATCTATGAGTTCCACAGAGGCTTC 1896

Qy 599 LeuLysGlnIleGlnGlnArgLeuAlaLeuTrpGlnLysArgSerAsnAlaGlnLysArg 618
Db 1897 CTGCGCGAGGTGAGACAGAGGCTGGACCTTGGAGAGGCGCTTCCAAAGCCACACAAA 1956
Qy 619 ---AspTrpGlnArgLysIleGlyAspValMetLeuLysAsnIleGlnLysMetLysIleLeu 637
Db 1957 GGCACTCATCAACGAATCGGGGACATCTGCTCAGGAACATGCGCCAGTTAAAGAGTTT 2016
Qy 638 AlaAlaHisLeuTrpLysHisSerGlnAlaLeuGlnAlaLeuGlnLysGlnLysSer 657
Db 2017 ACCAGCTACTTCCAAAGACATGACGAGGTCTTACAGAACTCGAAGAAAGCTTACCAAGC 2076
Qy 658 SerArgLysGlnGlnAsnPheCysArgAspPheGlnLeuGlnLysValCysTrpLeuPro 677
Db 2077 TGTAAAGATTGAGAGGCACTGATCAAGAGATTGAGTGCAGAAAGCTTGTCTACTGCT 2136
Qy 678 LeuAsnThrPheLeuLeuArgProLeuHisArgLysMetHisArgLysGlnValLeuGln 697
Db 2137 CTCAACACAGTTCTGCTGAAGCCCATCCAGCGGCTGCTGACCTACCGCTGCTGGCG 2196
Qy 698 ArgLeuCysLysHisAspProSerHisAlaAspPheArgAspCysArgAlaAlaLeu 717
Db 2197 CGCTTATGCGGACATTCAGGCCCGGCGACCATGACTACGCTGACTGCATGACGCGCTG 2256
Qy 718 AlaGlnIleThrGlnMetValAlaGlnLeuHisGlyThrMetIleLysMetGlnAsnPhe 737
Db 2257 AAAGCATCACAGAGATGACCAACACACTTACAGCACTTCTCATCTCGGCTGGAGAACTG 2316
Qy 738 GlnLysLeuHisGlnLysLeuLysAspLeuIleGlyIleAspAsnLeuValProGly 757
Db 2317 CAGAAAGCTAACGAGCTGACAGCGGACCTGTTGGGCATAGAAACCTCATTTGCTCTGGC 2376
Qy 758 ArgGlnPheIleArgLysGlySerLysSerLysLysSerGlyLysGlyLeuGlnGlnArg 777
Db 2377 AGGAGATTCACTCGTGAAGGCTGCTTTCACAAAGCTTACCAAGAAAGGCGCTCAGAGAGG 2436
Qy 778 MetPhePheLeuPheAsnAspValLeuLeuLysTrpHisSerArgGlyLeuThrAlaSerAsn 797
Db 2437 ATGTTTCTTCTGTTCTCAGATATGTGTGCTGACACAAAGAGATTGACAGGAGCCAGC 2496
Qy 798 GlnPheLysValHisGlnGlnLeuProLeuTrpGlyMetThrIleGlnGlnSerGlnAsp 817
Db 2497 CACTTCGCGAGTCGGGGGCTCTTCCCTCCAAAGCATGCTGATGGAAGAAAGATGATAC 2556
Qy 818 GlnTrpGlnValProHisCysLeuThrLeuArgLysGlnArgLysSerIleValAla 837
Db 2557 GAGTGTCTGTTCACACTGTTTCAACATTCACCGGCTCAGAAACATGATGATGCGCA 2616
Qy 838 AlaSerSerArgSerGlnMetGlnLysTrpValGlnAspIleGlnMetAlaIleAspLeu 857
Db 2617 GCCAGCACTCGGCTGAGAAAGAAAGATGATGCTGAGACCTCGGATCCGATCAAGCA 2676
Qy 858 AlaGlnLysSerSerSerProAlaProGlnPhe-----LeuAlaSerSerPro 873
Db 2677 GCCAAGAGTGGCGGTGACACGCGCCCTGCACTGCCAGGCGGCACATGTTGATCCTGCC 2736
Qy 874 ProAspAsnLysSerProAspGlnAlaThrAlaAlaAspGlnLysSerGlnAspLeu 893
Db 2737 CCC-----AGATCCCCCAACGAG---GATCTCTGAGACGAGAGTCAAGAAAGATGCT 2787
Qy 894 SerLysSerArgTrpThrSerLeuGlnArgLysAlaProHisArgGlyAsnThrMetValHis 913
Db 2788 CGGGGTGTCGCAACTCTCGAGAGGCGATGCGCAAGCCGAGGCAACCAACATGATGAC 2847
Qy 914 ValCysTrpHisArgAsnThrSerValSerMetValAspPheSerIleAlaValGlnAsn 933
Db 2848 GTGATGCTGATCCGGAACACACAGCTGTCCAGGCGACAGACCAAGTGCAGCTGTGAGAAC 2907
Qy 934 GlnLeuSerGlnAsnLeuLeuArgLysPheLysAsnSerGlnGlyTrpGlnLysLeuTrp 953
Db 2908 CAGCTTTCAGGATATCTGCTAAGAAAGTTCAAAAACAGTATGCTGCGCAGAAAGCTGCG 2967

Qy 1025 ArgSerAlaThrSerSerAlaSerArgProHISValLeuSerHisIleValSerLeuVal 1044
Db 1442 CGCAGTGCACACCTGCTGCTCGGACCCACGTTGTAGCCACAAAGATCTCTGTG 1501
Qy 1045 Tyr 1045
Db 1502 TAT 1504

RESULT 8
US-09-764-868-51
Sequence 51, Application US/09764868
Patent No. US2002016871A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PT232
CURRENT APPLICATION NUMBER: US/09/764,868
CURRENT FILING DATE: 2001-01-17
Prior application data removed - refer to PAM or file wrapper
NUMBER OF SEQ ID NOS: 1510
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 51
LENGTH: 1718
TYPE: DNA
ORGANISM: Homo sapiens
US-09-764-868-51

Alignment Scores:
Pred. No.: 2,66-86 Length: 1718
Score: 1006.50 Matches: 189
Percent Similarity: 78.7% Conservative: 47
Best Local Similarity: 63.0% Mismatches: 57
Query Match: 18.4% Indels: 7
Gaps: 3

US-09-555-342b-2 (1-1045) x US-09-764-868-51 (1-1718)

Qy 738 GlnIleValLeuHISGlnLeuValLeuValLeuValLeuValLeuValLeuValLeuVal 757
Db 13 CAGAAAGCTAACGAGAGCTGACGCGGACCTGGGCGCAAGAACTTCATTCCTGCTGCG 72
Qy 758 ArgGlnPheIleArgLeuValSerLeuSerValLeuSerValLeuSerValLeuSerVal 777
Db 73 AGGAGATTCATCCGTGAGGCTGCTTCACAAAGCTCAAGAAAGGCTTCAGAGAGG 132
Qy 778 MetPhePheLeuPheAspValLeuValLeuValLeuValLeuValLeuValLeuValLeu 797
Db 133 ATGTTTTCCTGTTCTCAGATATGTTGCTGTACCAAGCAAGAGATTGCGAGGACCGC 192
Qy 798 GlnPheValHISGlnLeuProLeuValLeuValLeuValLeuValLeuValLeuVal 817
Db 193 CACTTCGCGATCCGAGGCTCTCTCCCTCCCAAGGCAATGCTGAGAAAGATGATAC 252
Qy 818 GluTrpGlyValProHISValLeuValLeuValLeuValLeuValLeuValLeuValLeu 837
Db 253 GAGTGTCTGTTCCACACTGTTTCAACGATCAACGCGGCTGCAAAACATGCTGCTGCA 312
Qy 838 AlaSerSerArgSerGluMetGluValLeuValLeuValLeuValLeuValLeuValLeu 857
Db 313 GCCAGCACTCGGCTGAGAAAGAGATGATGCTGAGCTGAACTCCGCGATCCAGCA 372
Qy 858 AlaGluValSerSerSerProAlaProGluPhe-----LeuAlaSerSerPro 873
Db 373 GCCAAGAGTGGCGGTGACACGCGCTGACCTGACGAGCCGACACTGTGTGACCTGCTCC 432
Qy 874 ProAspValLeuSerProAspGluValLeuValLeuValLeuValLeuValLeuValLeu 893
Db 433 CCC-----AGATCCCAACAGAG---GTATCTCTGAGCGAGGATCGAAGATGATGCT 483
Qy 894 SerAlaSerArgTrpSerLeuGluArgGlnAlaProHISArgGlyAsnTrpMetValHIS 913
Db 484 CGGGGTGTCCGACACTCTCTGAGAGGGGATGCGCAGCAGCGGGCAACACCAATGAC 543

Qy 914 ValCysTrpHISArgAsnThrSerValSerMetValAspPheSerIleAlaValGluAsn 933
Db 544 GTGTGCTGTGATCCGAAACACGAGGTGTCCAGGCGAGACCAAGGCACTGTGTGAGAAC 603
Qy 934 GlnLeuSerGlyAsnLeuValLeuValLeuValLeuValLeuValLeuValLeuValLeu 953
Db 604 CAGCTTCAGAGATATCTGCTAAGAAAGTTCAAAACAGTATGCTGTGCAAACTCTGCG 663
Qy 954 ValValPheTrpAspPheCysLeuPhePheTrpLeuSerHisGlnAspAsnHISProLeu 973
Db 664 GTCCCTTTACCACTCTGTTGTTCTTCAAAACTCATTCAGATGATCTACCACTG 723
Qy 974 AlaSerLeuProLeuValLeuValLeuValLeuValLeuValLeuValLeuValLeu 993
Db 724 GCCAGCTCCCGCTGCTGAGGTGATGAGCATCCCAAGGAGGCCGATGCGATCAC 783
Qy 994 LysAspTrpValPheValLeuHISPheLeuSerHisValTrpTrpPheArgAlaGluSer 1013
Db 784 AAAGACTATGTTTCAAGCTCAGTTCAATCCCAAGCTTACTTCTTCGCGGCTGAGAGC 843
Qy 1014 GluTrpTrpPheGluArgTrpMetGluValIleArgSerAlaThrSerSerAlaSerArg 1033
Db 844 AAGTACACATTGGAAGGTGATGAGGTGATCCAGGGGCGCAGCAGCTCAGCCGGAGG 903

RESULT 9
US-11-097-143-3041
Sequence 3041, Application US/11097143
Publication No. US2005020858A1
GENERAL INFORMATION:
APPLICANT: Venter, J. Craig
APPLICANT: et al.
TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
FILE REFERENCE: CL000728
CURRENT APPLICATION NUMBER: US/11/097,143
PRIOR FILING DATE: 2005-04-06
PRIOR APPLICATION NUMBER: 60/157,832
PRIOR FILING DATE: 1999-10-05
PRIOR APPLICATION NUMBER: 60/160,191
PRIOR FILING DATE: 1999-10-19
PRIOR APPLICATION NUMBER: 60/161,932
PRIOR FILING DATE: 1999-10-28
PRIOR APPLICATION NUMBER: 60/164,769
PRIOR FILING DATE: 1999-11-12
PRIOR APPLICATION NUMBER: 60/173,383
PRIOR FILING DATE: 1999-12-28
PRIOR APPLICATION NUMBER: 60/175,693
PRIOR FILING DATE: 2000-01-12
PRIOR APPLICATION NUMBER: 60/184,831
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: 60/191,637
PRIOR FILING DATE: 2000-03-23
NUMBER OF SEQ ID NOS: 43008
SOFTWARE: ParseSeq for Windows Version 4.0
SEQ ID NO 3041
LENGTH: 1747
TYPE: DNA
ORGANISM: DROSOPHILA
US-11-097-143-3041

Alignment Scores:
Pred. No.: 2,75e-77 Length: 1747
Score: 913.50 Matches: 217
Percent Similarity: 54.0% Conservative: 110
Best Local Similarity: 35.8% Mismatches: 207
Query Match: 16.7% Indels: 73
Gaps: 14

US-09-555-342b-2 (1-1045) x US-11-097-143-3041 (1-1747)

Qy 423 GlySerHisProSerProAlaProArgArgSerProAlaGluAsnValGlnAlaAspGly 442

PRIOR APPLICATION NUMBER: 60/175,693
 PRIOR FILING DATE: 2000-01-12
 PRIOR APPLICATION NUMBER: 60/184,831
 PRIOR FILING DATE: 2000-02-24
 PRIOR APPLICATION NUMBER: 60/191,637
 PRIOR FILING DATE: 2000-03-23
 NUMBER OF SEQ ID NOS: 43008
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO: 743
 LENGTH: 1220
 TYPE: DNA
 ORGANISM: DROSOPHILA
 US-11-097-143-743

Alignment Scores:
 Pred. No.: 9.4e-77 Length: 1220
 Score: 905.50 Matches: 180
 Percent Similarity: 67.84 Conservative: 69
 Best Local Similarity: 49.04 Mismatches: 105
 Query Match: 16.64 Indels: 13
 DB: 13 Gaps: 7

US-09-555-342B-2 (1-1045) x US-11-097-143-743 (1-1220)

QY 8 ProthrProglySerArgLeuGluValAlaProGluAaSerGlyIleSerThrLeuGluArg 27
 DB 102 CCAAGGGGGGGGGCTGCTGCGGAGGCGGATGAGCACTCCCTCAACACTCTTCGCGG 161
 QY 28 GlyInLysProProProProThrPro-----SerGlyLysLeuValSerIleLysIle 44
 DB 162 GTGATGAGCAGCCCTCCACACCCGACATCGCGGCGGAAAAAGCTGACTGTTCGATC 221
 QY 45 GlnMetLeuAaPheThrGlnGluValAlaPheGluValProGlnArgAlaProGlyLysVal 64
 DB 222 CAGATGCTGAGCAGCTCATCATCCATGTTCCAAAGTACAGGCTTAAAGCACCTGGCGCTG 281
 QY 65 LeuLeuAaPheValIleCysAaSerIleLeuAaValGluGlyAaPheThrPheGlyLeuGlu 84
 DB 282 CTGTTTGAACAGTATGCGGCTGATGATTTACTGAGGCGGCACTTCTGCGCTTGAG 341
 QY 85 PheProAaP-----HisLysLysIleThrValITPLeuAaPheLeuLysProIle 101
 DB 342 TACAGGAAGTCTCCACACATACCAA-----TACTGGCTGATCTGAGAAAGCAATG 395
 QY 102 ValLysGlnIleArgArgProLys---HisValValValLysPheValValLysPhe 120
 DB 396 AACGGCCAGGGGCTCTCTTATCATTCAGTCCAGCTGCGCTTGCATCAAGTTCTAT 455
 QY 121 ProProAaPheIleThrGlnLeuGlnGluGluLeuThrArgTyrLeuPheAlaLeuGlnVal 140
 DB 456 ACACCCGATCCAGGCCAATTGGAAGAGATACACCAAGTATTTTCTGCTGCAATC 515
 QY 141 LysGlnAaPheLysAlaGlnGlyArgLeuThrCysAaAaPheThrSerAlaAlaLeuLeuIle 160
 DB 516 AACGAGACCTGGCCACAGGCAAGCTGCGAGTGCACAGTAAACACGCGGCTTAAATGCA 575
 QY 161 SerHisIleValGlnSerGluIleGlyAaPhe-----AaGluAlaLeuAaPheArgIle 178
 DB 576 AGCTACATTTGACAGGCTCATGCGCGCATCTGCTCCGAGGATTAATCCGACCATACT 635
 QY 179 HisLeuAlaLysAaLysLysTyrIleProGlnGlnAaPheAa-----LeuGluAaPheLysIleVal 197
 DB 636 TATCTGCTCTGATCGCTTGTGTGCGCAACAGGAGCGCACCATCATCAGCGAAAGTTATG 695
 QY 198 GluPheIleHisAaSerHisIleGlyGlnThrProAlaGlnLysAaPheGlnLeuLeuGlu 217
 DB 696 GAGAAACCAAAACATGTTGTGTCATCCCGGCGGAGCGGACCTTAACCTTTGAG 755
 QY 218 IleAlaAaArgLeuGluMetCysTyrIleArgLeuHisPheProAlaLysAaPheArgIle 237
 DB 756 ACGCTCGAGAGTGTGAGCTGATGCGATGAAGAAAGCATCCGCGCAAGAGATGTGAAGG 815
 QY 238 ThrLysIleAaPheLysValAlaAlaAaThrGlyIleLeuValPheGlnGlyPheThrLys 257

DB 816 GTGGCGTTAATCTGCTGTGGCCACATGGGCATCACAGTCTTCCAAATCATCGCGG 875
 QY 258 IleAaAlaPheAaSerIleAlaLysValIleGlyLysLeuSerPheLysArgLysPheLeu 277
 DB 876 ATCAACACCTTCTCGTGGGCTTAAGATGACAAAGATTTCTTCAAGCGCAAGGATTCCTG 935
 QY 278 IleLysLeuAaArgProAaPheAlaAaSerAlaTyrGlnAaPheThrLeuGluPheLeuVal 297
 DB 936 GTCAACTGCATCCGAGAGATATGATTAACAAGATACCGTAGATTCTTCTTCGAG 995
 QY 298 SerArgAaPheCysLysSerPheThrLysIleCysValGlnHisAlaPhePheArg 317
 DB 996 GGTGCAACGAGTGCAGAAACCTTGGAAAAAATGCGTGAATAATCAGGATTCCTCGA 1055
 QY 318 LeuPheGlu---GluProLysPheProLysPheProValLeuPheSerThrGlySer 336
 DB 1056 TGCACCTGCCGTAACAAATACGCCGCAAAATCGCGTCTCTCGCGGGTAGTTCA 1115
 QY 337 PheArgPheSerGlyArgThrGlnLysGlnValLeuAaPheTyrValLysGluGlyHis 356
 DB 1116 TTCGCTATAGCGGAAAAACCGAGAGCATTAATGAGTTGTTGCGGAAAAATTATGTG 1175
 QY 357 LysLysValGlnPheGluArg 363
 DB 1176 AAGCCCAAAACTTCCAAAG 1196

RESULT 11

US-11-097-143-3040/c
 Sequence 3040, Application US/11097143
 Publication No. US2005020858A1

GENERAL INFORMATION:

APPLICANT: Venter, J. Craig
 APPLICANT: et al.
 TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
 TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
 FILE REFERENCE: CLO00728

CURRENT APPLICATION NUMBER: US/11/097,143

PRIOR FILING DATE: 2005-04-04

PRIOR APPLICATION NUMBER: 60/157,832

PRIOR FILING DATE: 1999-10-05

PRIOR APPLICATION NUMBER: 60/160,191

PRIOR FILING DATE: 1999-10-19

PRIOR APPLICATION NUMBER: 60/161,932

PRIOR FILING DATE: 1999-10-28

PRIOR APPLICATION NUMBER: 60/164,769

PRIOR FILING DATE: 1999-11-12

PRIOR APPLICATION NUMBER: 60/173,383

PRIOR FILING DATE: 1999-12-28

PRIOR APPLICATION NUMBER: 60/175,693

PRIOR FILING DATE: 2000-01-12

PRIOR APPLICATION NUMBER: 60/184,831

PRIOR FILING DATE: 2000-02-24

PRIOR APPLICATION NUMBER: 60/191,637

PRIOR FILING DATE: 2000-03-23

NUMBER OF SEQ ID NOS: 43008

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 3040

LENGTH: 4150

TYPE: DNA

ORGANISM: DROSOPHILA

US-11-097-143-3040

Alignment Scores:

Pred. No.: 1.42e-74 Length: 4150
 Score: 891.50 Matches: 222
 Percent Similarity: 45.4% Conservative: 113
 Best Local Similarity: 30.1% Mismatches: 201
 Query Match: 16.3% Indels: 202
 DB: 13 Gaps: 15

US-09-555-342B-2 (1-1045) x US-11-097-143-3040 (1-4150)

QY 423 GYSerHisProSerProAlaProArgArgSerProAlaGlyAenlySGlnAlaAspGly 442
 DB 3064 GGATCGAGTAAAGTCTTGTCTCACTCCGATCTGAGCGAGAAACA----- 3020
 QY 443 AlaAlaSerAlaProThrGluGluGluGluValValValysAspArgThrGlnGlnSer 462
 DB 3019 -----TATGACGTAAATACAGGAGAAATCCACACGCT 2987
 QY 463 LysProGlnProProGlnProSerThrGlySerLeuThrGlySerProHisLeuSerGlu 482
 DB 2986 AAGGAGAG-----GAGCTAGCCACAGCGCTGCCACCGCTCG----- 2948
 QY 483 LeuSerValAsnSerGlnGlyValAlaProAlaAsn---ValThrLeuSerProAsn 501
 DB 2947 -----CACCGAATGGAGACGGAACCGAAATGGGCAATCCCTGAGC----- 2906
 QY 502 LeuSerProAspThrLysGlnAlaSerProLeuHisSerProLeuLeuAsnAspGlnAla 521
 DB 2906 ----- 2906
 QY 522 CysProArgThrAspAspGluAspGluGlyArgArgLysAspRheProThrAspLysAla 541
 DB 2905 ---ACACAAACAGATATTGAGGAGCAAGTGAAGGAAATGGCCACCGAGCCGAGC 2849
 QY 542 TyrPheIleAlaLysGluValSerThrThrGluArgThrTyrLeuLysAspLeuGluVal 561
 DB 2848 TACTTTTGGCCAAAGAGCTGTGATGACGGAACGTACGTAACAAAGAGATCTAGATGTG 2789
 QY 562 IleThrSerTrpPheGlnSerThrValSerLysGluAspAlaMetProGluAlaLeuLys 581
 DB 2788 CTGAACCAACACTTCCAGAGGTGTGATGCTCGGAGATGTG-----GAGCAAGCTCGAG 2735
 QY 582 SerLeuIlePheProAsnPheGluProLeuHisLysPheHisThrAsnPheLeuGlu 601
 DB 2734 CCGGTG---TTGAGCTGCTGAGCTCCCTGGCCGAGACCAATCTCTTCCTGCGGAC 2678
 QY 602 IleGluGlnArgLeuAlaLeuTrpGluGlyArgSerAsnAlaGlnIleArgAspTyrGln 621
 DB 2677 ATCGAACACCGCATGCTGAGCTGAGGAGGAGAGAGAGG-----CACGAAGCCGAC 2627
 QY 622 ArgIleGlyAspValMetLeuLysAsnIleGlnGlyMetLysHisLeuAlaHisLeu 641
 DB 2626 CGCATCGGAAGACGTATGATGAACACATGCGCCCTGCTGCCATCTGATGAGTACGTG 2567
 QY 642 TrpLysHisSerGluAlaLeuGluAlaLeuGluAsnGlyLysSerSerAspArgLeu 661
 DB 2566 CAGACGACCTGACATCTTGCACTGTATGAACGATGACGAAGCGATGAACGCTTC 2507
 QY 662 GluAsnPheCysAspAspPheGluLeuGlnLysValCysTyrLeuProLeuAsnThrPhe 681
 DB 2506 CGTCAGGTGTAACAAGAGATTGAGCAAGAAAGGTTTGCTATCAACCATCGGAACTT 2447
 QY 682 LeuLeuArgProLeuHisArgLeuMetHisTyrLysGlnValLeuGluArgLeuCysLys 701
 DB 2446 CTACTGAGCCCTCAACCGCTGTGCACTACCACTGATCTTAAGCGCGCTGTGCGAC 2387
 QY 702 HisHisProProSerHisAlaAspPheArgAspArgAlaAlaLeuAlaGlnIleThr 721
 DB 2386 TACTATGGGAGAGACATATGCACTATGCGATGCGATGCGTACGCAACTTGTCTGTT 2327
 QY 722 GluMetValAlaGlnLeuHisGlyThrMetIleLysMetGluAsnPheGlnLysLeuHis 741
 DB 2326 CGCAGACCAAGAGTATTAGGTCCCAAGCTTCCGACCTCTGCAAACTTTGAGAGCTGTGT 2267
 QY 742 GluLeuLysLysAspLeuIleGlyIleAspAsnLeuValProGlyArgGluPheIle 761
 DB 2266 GAATCGAAGCGGAC---ATCAACTTCAGAGCATGTGTTCCAGCCCACTCCGCGCTATTC 2210
 QY 762 ArgLeuGlySerLeuSerLysLeuSerGlyLysGlyLeuGlnGlnAspMetPhePheLeu 781
 DB 2209 CGCCAGGAGATGCTTCTGAAAGCACTCAAGCGCGGTCTGACGACGAGAGATGTCTTCTTG 2150

QY 782 PheAsnAspValLeuLeuTyrThrSerArgGlyLeuThrAlaSerAsnGlnPheLysVal 801
 DB 2149 TTCTCCGACCTGCTGCTCTAAGGTGAAGTCC---CCGCTGAGCAAGAGCTTCCGATTT 2093
 QY 802 HisGlyGlnLeuProLeuTyrGlyMetThrIleGluGluSerGluAspGluTyrGlyVal 821
 DB 2092 TTAGGCCACGCTCTGCTGCTTCACTACCGAGAAACCCCGGAC----- 2045
 QY 822 ProHisCysLeuThrLeuAspGlyGlnArgGlnSerIleIleValAlaAlaSerSerArg 841
 DB 2044 ---AACACTTCTCACTTTCGGGAGACAGTGTGCATCAACAGTGAAGTGGGACACACC 1968
 QY 842 SerGluMetGluLysTrpValGluAspIle---GlnMetAlaIleAspLeu----- 857
 DB 1987 GGTGAAAAGACCTCTGCTGCTGGCGAGCTCCAGAGCGCGCGGATATTAAAAACGA 1928
 QY 857 ----- 857
 DB 1927 CCGGCCAACATGCACTCCAGCTACGACGCTCAAGAACTGCAGTAGTATATATTG 1868
 QY 857 ----- 857
 DB 1867 TATTAATGCGATTTCGATGCTTTGTTTCTTAATGTAGCGTTTCAACAAAGTCCCATATA 1808
 QY 857 ----- 857
 DB 1807 TGAATTGAATAGCTTTAAGAACCAAGCTACCGGAATGTAAATATCCAAATAACTCTT 1748
 QY 857 ----- 857
 DB 1747 ATAAACCAAAAGCTTTGGTGCAGCGGCTAACTATTACCTCAACAAATAAAAAGTTT 1668
 QY 857 ----- 857
 DB 1687 AACTTAACTAATCTTTTGTAACTCCCTGTTTCTCAATTTGTTTGTGCTGGC 1628
 QY 858 ----- 1628
 DB 1627 AAGAAAAGACGATGATGCGCGCGTAAAGAAAGATGACCAACAAATCCAAATCTT 1568
 QY 869 ----- 869
 DB 1567 CTTCGAACAAACCCAGTAATGTATTTCACCTTAATTTTCAGGCTCCGAGAGAGG 1508
 QY 870 -----AlaSerSerProAspAsnLysSerProAspGluIleThrAlaAla 885
 DB 1507 CTTGACCTATTGGGTCTGAGCAACGGAACAACAGAGCTTGAACGAGTGTGAACGG 1448
 QY 886 AspGlnGluSerGlu-----AspAspLeuSerAlaSerArgThrSerLeuGluArg 902
 DB 1447 CGCGGGTCCGCTGACTACGACGAGCAAGAGTTGCAAGTTCAGCTCAGACACAGCAAAACAG 1388
 QY 903 GlnAla-ProHisArgGlyAsnThrMetValHisValCysTrpHisArgAsnThrSerVal 922
 DB 1387 GAGCGAGCCATCCCGAAGCAATACGCGTCTGACAGTCTGTGGCATCTGTGCGCCACCGT 1328
 QY 922 LserMetValAspPheSerIleAlaValGluAsnGlnLeuSerGlyAsnLeuAspGly 942
 DB 1327 AGACATGGGAGATCACTGATTAAGCCCGCCAGACCAAGTTCGCGGTATATGCTGCGAAA 1268
 QY 942 sPheLysAsnSerAsnGlyTrpGlnLysLeuTrpValPheThrAsnPheCysLeuPhe 962
 DB 1267 GTTCAAGAAATAGCTCCGCTGCGAGAAAGCTCGGTGTGTTCACGCTTTTGTCTGTA 1208
 QY 962 ePheTyrLysSerHisGlnAspAsnHisProLeuAlaSerLeuProLeuLeuGlyTyrSe 982
 DB 1207 CTTTTCAAAGAGCTACAGGACAGAGTTCGCACTGGCTAGTCTCCGCTATATGGTTACAC 1148
 QY 982 rLeuThrIleProSerGluSerGluAsnIleGlnLysAspTyrValPheLysLeuHisPhe 1002
 DB 1147 GGTGGGTCTCTCTGTCACCAAGATGCGCTTCAGAAAGAGTGTGCTTCAAGCTTCTCTT 1088
 QY 1002 eLysSerHisValTyrTyrPheArgAlaGluSerGluTyrThrPheGluArg 1019

Db 1087 TAAAGACCACTGCTTCTTCCGCGCGAAGTCCGACACCTTACAAAG 1036
RESULT 12
US-10-755-889-499
Sequence 499, Application US/10755889
Publication No. US20040171823A1
GENERAL INFORMATION:
APPLICANT: Bristol-Myers Squibb Company
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES ASSOCIATED WITH THE NF-kB
FILE REFERENCE: D0284 NP
CURRENT APPLICATION NUMBER: US/10/755,889
PRIORITY FILING DATE: 2004-01-13
PRIORITY FILING DATE: 2003-01-14
PRIORITY FILING DATE: 2003-01-14
PRIORITY FILING DATE: 2003-05-12
NUMBER OF SEQ ID NOS: 823
SOFTWARE: PatentIn version 3.2
SEQ ID NO 499
LENGTH: 3309
TYPE: DNA
ORGANISM: Homo sapiens
US-10-755-889-499
Alignment Scores:
Pred. No.: 7,2e-61 Length: 3309
Score: 748.50 Matches: 195
Percent Similarity: 49.6% Conservative: 94
Best Local Similarity: 33.4% Mismatches: 233
Query Match: 13.7% Indels: 61
Gaps: 15
US-09-555-342b-2 (1-1045) x US-10-755-889-499 (1-3309)
Qy 18 GluHensSerGlyIleSerThrLeuGluArgGly-----GlnIysProPro 32
Db 331 CAGAAATATCATGAGAACTCTCTCGGTCCATTAAAGATTGTCMAAAAGCTT--- 387
Qy 33 ProThrProSerGlyIysLeuValSerIleIysIleGlnMetLeuAspThrGlnGlu 52
Db 388 -----AAAGCATGCAGTGCMAAGTGAATCTTCTGATGATGATCAGAAATAT 432
Qy 53 AlaPheGluValProGlnArgAlaProGlyIysValLeuLeuAspAlaValCysAsnHis 72
Db 433 ACCTGTGTGTAGAGAAACGCTCCAGAGACAAAGTGTGTGATTAAGTGTGAACAC 492
Qy 73 LeuHnLeuValGluGlyAspTyrPheGlyLeuGluPheProAspHisIleIysValThr 92
Db 493 TTGAATCTTGTAGAGAAAGACTACTTGTGGCTTACGTATCCAGATGCTGAACACAGAG 552
Qy 93 ValIlePheAspLeuLeuIysProIleValIysGlnIleArgArgProIys---HisVal 111
Db 553 AATGTGTGAGACCTGCTAAGAAATTAATAAAACAGGTTCGAAGTGTGCTGGCAC--- 609
Qy 112 ValValIysPheValValIysPhePheProProAspHisIleGlnLeuGlnGluLeu 131
Db 610 ---TTTCATTATTAAATTTATATCCACCAAGCCTGCCCACTATCTGAAGATATC 666
Qy 132 ThrArgTyrLeuPheAlaLeuGlnValIysGlnAspLeuAlaGlnGlyArgLeuThrCys 151
Db 667 ACCAGGATCTACCTGCTTGTGAGTGGAGATGCATGCTGTCCGAAAGGCTGCCCTGC 726
Qy 152 AsnAspThrSerAlaAlaLeuLeuIleSerHisIleValGlnSerGlnIleGlyAspPhe 171
Db 727 TCCTTTGTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 786
Qy 172 Asp---GluAlaLeuAspArgGlnIleIleValIysValIysValIysValIysValIys 189
Db 787 GACCCAGATGAATGTGGAGAGATTAATGATGATGCTTCCGCTTGGACCAACCAACACT 846
Qy 190 AspAlaLeuGluAspArgIleValGluPheHisIleAsnHisIleGlyGlnThrProAla 209

Db 847 AAAGAACTGGAAGCAAAAGTATGAGCTGACCAAGACCAAGACCAAGAAATGACCGCACCA 906
Qy 210 GluSerAspPheGlnLeuGlnIleAlaArgArgLeuGlnMetTyrGlyIleArgLeu 229
Db 907 GAAGCAGAGATGCATTTCTTGAAATAATCCAAAAATTAATCAATGATGAGGATGATTTA 966
Qy 230 HisProAlaIysAspArgGluGlyThrIysIleAsnIleValAlaAsnThrGlyIle 249
Db 967 CATCATGCTTAAGACTCAAGAGGGGTAGAAATTAATGATTAAGAGTTTGTGCAGAGTGTCTG 1026
Qy 250 LeuValPheGlnGlyPheThrIysIleAsnIlePheAsnProAlaIysValArgIysLeu 269
Db 1027 TTGAATATATCGGACCGGCTCGAATAAAGATTTGCTGCGCCCAAGCTTCAAAAGATT 1086
Qy 270 SerPheIysValGlyValArgPheLeuIleIysLeuAspArgProAspAlaAsnSerAlaTyrGln 289
Db 1087 TCATCAAAACGGAACAACTTTTACATTAAGATCCGCGCGAGAGATTGAACAAATTTGAA 1146
Qy 290 AspThrLeuGluPheLeuMetAlaSerArgAspPheCysIysSerPheThrIysIleCys 309
Db 1147 AGCACTATGGGATTAAAGCTGCCAAACATGACGCTGCCAAGCTTTATGAAATGATATG 1206
Qy 310 ValGlnHisIleAlaPhePheArgLeuPheGlnGluIysProIysProIysProIysProVal 329
Db 1207 GTTGAAGATCATATCTTTTCAAGACTAG---TTACCAAGAACACTCCCAAGAAATTC 1263
Qy 330 LeuPheSerArgIysSerSerPheArgPheSerGlyArgThrGlnIysGlnValLeuAsp 349
Db 1264 CTA---ACCTTGGGTTCMAATGTTTCGTTATGATGTCAGAGAACACAAAGCCAAACGAAAGA 1320
Qy 350 TyrValIysGluGlyGlnIleIysValIysValGlnPheGluArgIysIleSerIysIleHis 369
Db 1321 GCCAGCTGCTGATTAAGATGCCGCCACACTTAATTGAACCTGATCCAGCAAAAGCTTAT 1380
Qy 370 SerIle---ArgSerLeuAlaSerGlnProThrGlnLeuAsnSerGluVal---LeuGlu 387
Db 1381 ACCATGTCCTCAGAGTTGATGAGACATGACATGAAATCAATGAATATATCATGAAG 1440
Qy 388 GlnSerGlnGlnSerThrSerLeuThrPheGlyGlnGlyAlaGluSerProGly----- 405
Db 1441 GATTCTATGTCGTGCGACAGAGTTGTACTGCGCAGTACGCGCACAAACAAAGGCACTCTCT 1500
Qy 406 -----GlyGlnSerCysArgArgGly 412
Db 1501 CAGACCAACTGATCACCACTGTGACTCCGAGAAAGAGCTGAGAGAGAGCGGAGCAG 1560
Qy 413 LysGluProIysValSerAlaGlyIysProGlySerHisProSerProAlaProArgArg 432
Db 1561 GAAGAGACAAACGAGAGAGGGGAGAGAAATCAACGCCCATCTGCGCATCCAGACAGAG 1620
Qy 433 SerProAlaGlyAsnIysGlnAlaAspArgIleAlaSerAlaProThrGlnGlnGluGlu 452
Db 1621 GGAAGACTGACAGAGGACGACGACCCAGCGCAGCGGAGAGACCACTGCCACTAG 1680
Qy 453 GluValValIysAspArgThrGlnGlnSerIysProGlnProProGlnProSerThrGly 472
Db 1681 GAGCTGAAATAAATCAAGATGACTGATGAACATCAA----- 1719
Qy 473 SerLeuThrIysSerProHisIleLeuSerGluLeuSerValAsnSerGlnGlyValAla 492
Db 1720 -----ACCAACATTAGAGACTGAAGAAACCTTTTGAACACTCAACAGACACTGCC 1773
Qy 493 ProAlaAsn-----ValThrIysSerProAsnLeuSerProAspThrIys 507
Db 1774 GTAACGAATGAATGAGAGAAAGAGCTTCCACTCCCGGTGCACTGCGCGCCAGGAG 1833
Qy 508 GlnAlaSerProLeuIleSerProLeuLeuAsnArgIleAlaCysProArgThrAspAsp 527
Db 1834 GAGATGCCCCCAATGATGAACACTTGTCCCTGAAGAGAAATG---GAAACCAAGACG 1890
Qy 528 GluAspArgIysArgArgIysArgPheProThrAspIysAlaTyrPheIleAlaIysGlu 547

Db 1891 GAGTCACGTGATAGAGACGACCCACCG----- 1920
Qy 548 ValSerThrThrGluArgThrTyrLeuLysAspLeuGluValIleThrSerTyrPheGln 567
Db 1921 ---TGCACCCACCTCCGCTTAGCAGTGAAGGTGTGCAGAGAACCGTGTGTGGAGG 1977
Qy 568 SerThrVal 570
Db 1978 AGCCGGCGTG 1986

RESULT 13
US-10-956-157-1800
; Sequence 1800, Application US/10956157
; Publication No. US20050118625A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mount, William
; TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH
; TITLE OF INVENTION: HUMAN OSTEOARTHRITIS AND HUMAN PROTEASES
; FILE REFERENCE: 031896-043000 (AM 101081)
; CURRENT APPLICATION NUMBER: US/10/956,157
; NUMBER OF FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 319805
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1800
; LENGTH: 3309
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-956-157-1800

Alignment Scores:
Pred. No.: 7,2e-61 Length: 3309
Score: 748.50 Matches: 195
Percent Similarity: 49.6% Conservative: 94
Best Local Similarity: 33.4% Mismatches: 233
Query Match: 10.7% Indels: 61
Gaps: 15

US-09-555-342b-2 (1-1045) x US-10-956-157-1800 (1-3309)
Qy 18 GluAsnSerGlyIleSerThrLeuGluArgGly-----GlnLysProPro 32
Db 331 CAGAAATCATCTACAGTAAACTCTCGCTCCATTAAAGATTGCAGAAAACCT--- 387
Qy 33 ProThrProSerGlyLysLeuValSerIleValIleGlnMetLeuAspAspThrGlnGlu 52
Db 388 -----AAAGCATGCAGTGCAGAGTGATCTTCGATGATCGAATAT 432
Qy 53 AlaPheGluValProGlnArgAlaProGlyLysValIleLeuAspAlaValCysAsnHis 72
Db 433 ACCGTGATGTAGAGAAACGCTCCAGAGCAAGTGGCTTGTATTAAGTGTGTGAACAC 492
Qy 73 LeuAsnLeuValGluGlyAspTyrPheGlyLeuGluPheProAspHisLysLysIleThr 92
Db 493 TTGAACCTGTGAGAGAAAGACTACTTTGGGCTTACGATACGATGCTGAAACACAGAG 552
Qy 93 ValIlePheAspLeuLysProIleValIleGlnIleArgArgProLys---HisVal 111
Db 553 AATTGGTTGGACCTCGTAAAGAAATATAAAACAGGTTCGAAAGTGTCTGGCAC--- 609
Qy 112 ValValLysPheValValLysPhePheProAspHisThrGlnLeuGlnGluLeu 131
Db 610 ---TTTTCATTATATGTGAATTTTATTCACACAGACCTGCCAATATTCGAAAGATATC 666
Qy 132 ThrArgTyrLeuPheAlaLeuGlnValLysGlnAspLeuAlaGlnGlyArgLeuThrCys 151
Db 667 ACCGAGTACTACCTCTGCTGCAGTGTGCAGATGACATCGTGTCCGAGAGGCTGCCCTGC 726
Qy 152 AsnAspThrSerAlaAlaLeuLeuIleSerHisIleValGlnSerGlnIleGlyAspPhe 171
Db 727 TCTCTTGTATCCCTGGCGCTGTGGCTCTTACCTGCTCAGTGTGACAGAGCTGAGACTAT 786
Qy 172 Asp---GluAlaLeuAspArgGlnHisLeuValLysValLysValTyrIleProGlnGln--- 189

Db 787 GACCCAGATGATATCTGGAGAGATTAATTAGTAGTTCGCTTGCACCAACACACT 846
Qy 190 AspAlaLeuGluAspLysIleValGlnPheHisIleAsnHisIleGlyGlnThrProAla 209
Db 847 AAGAACTTGAGAAAGCAAGATGATGAGCTGCACAAAGGCCACAGAGAAATGACCCAGCA 906
Qy 210 GluSerAspPheGlnLeuLeuGluIleAlaArgArgLeuGlnMetTyrGlyIleArgLeu 229
Db 907 GAGCAGAGATGCATTTCTTGAAATATCCAAAATATATCAATGATATGGGAGATTATTA 966
Qy 230 HisProAlaLysAspArgGluGlyThrLysIleAsnLeuAlaValAlaAsnThrGlyIle 249
Db 967 CATCATGCTAAGGAGCTCAGAAAGGGGTGAATATGTATGAGAGTTTGGCAAGTGCTGTG 1026
Qy 250 LeuValPheGlnGlyPheThrLysIleAsnAlaPheAsnThrAlaLysValArgLysLeu 269
Db 1027 TTGATATATCGCAGCCGCTGCGAATTAACAGATTTTCCCTGGCCCAAGGTTCTTAAAGATT 1086
Qy 270 SerPheLysArgLysArgPheLeuIleLysLeuArgProAspAlaAsnSerAlaTyrGln 289
Db 1087 TCATACAAACGAAACAATTTTACATTATACGTCGCGCGGAGAGTTTGAACAATTGAA 1146
Qy 290 AspThrLeuGluPheLeuMetAlaSerArgAspPheCysLysSerPheTyrLysIleCys 309
Db 1147 AGCAGCATGGGTTTAACTGCCAACAACATCGAGCTGCCAAGCGTTATGGAAGATATGT 1206
Qy 310 ValGlnHisIleAlaPhePheArgLeuPheGluGluProLysProLysProVal 329
Db 1207 GTTGAGCATCATCATTTTTCAGCTACTG---TTACCAAGACCACTCCCAAAAATTC 1263
Qy 330 LeuPheSerArgLysSerSerPheArgPheSerGlyArgThrGlnLysGlnValLeuAsp 349
Db 1264 CTA---ACCTTGGTTTCAAGTTTCGTTATATGTGGCAGACACAGCGCAAGAGAA 1320
Qy 350 TyrValLysGluGlyGlyHisLysLysValGlnPheGluArgLysHisSerLysIleHis 369
Db 1321 GCCAGTCGTGTATAGATCCGCCAGCACCTTACTTTCAGCTCATCGACAAACGTTAT 1380
Qy 370 SerIle---ArgSerLeuAlaSerGlnProThrGluLeuAsnSerGluVal---LeuGlu 387
Db 1381 ACCATGCTCTGCGCTTGATGAGATCAGTGAATGAAACCATGAAATATATCATGAG 1440
Qy 388 GluSerGlnGlnSerThrSerLeuThrPheGlyGluGlyAlaGlnSerProGly----- 405
Db 1441 GATTCTATGTCTGCTGCAAGGTGTGTACTGCGCAGTACGCCCAACAAAAGGATCTCT 1500
Qy 406 -----GlyGlnSerCysArgArgGly 412
Db 1501 CAGACCAACTTGATCACCACTGTGACTCCGAGAGAAAGGCTGAGAGAGCGGAGACAG 1560
Qy 413 LysGluProLysValSerIleArgLysIleProGlySerHisIleProSerProAlaProArg 432
Db 1561 GAAAGGACAAACGAGAGAGGGGAGAAAGTCAAGCCCATCTTCGCGCATCCACAGAG 1620
Qy 433 SerProAlaGlyAsnLysGlnAlaAspGlyAlaAlaSerAlaProThrGlnGluGlu 452
Db 1621 GGAAGACTACAGTACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 1680
Qy 453 GluValValLysAspArgThrGlnGlnSerLysProGlnProProGlnProSerThrGly 472
Db 1681 GAGCTAGAAAACCTCAAGATGACCTGATGAACATCA----- 1719
Qy 473 SerLeuThrGlySerProHisLeuSerGlnLysSerValAsnSerGlnGlyValAla 492
Db 1720 -----ACCAAACTTACGAGCTGAAGAAACCTTCTTAAACCTTCAACAGACATGCC 1773
Qy 493 ProAlaAsn-----ValThrLeuSerProAsnLeuSerProAspThrLys 507
Db 1774 GTAACGAATATAGGAGAGAGGCTTTCACCTCCCGCTGGAGCTGGCGGACAGAG 1833
Qy 508 GlnLaserProLeuIleSerProLeuLeuAsnAspGlnAlaCysProArgThrAspAsp 527
Db 527

```

Db      1834 GAGATGCCCCCATATATGAAACCACTTGTCTCTGGAAGAAAAATG---GAAACCAAGAC 1890
Qy      528 GlnuepGluGlyArgArgLyseArgPheProThraePlySaIaTyPheIleAlaYseGlu 547
Db      1891 GAGTCCAGTGGATTAAGACGGAAACCCACCG----- 1922
Qy      548 ValSerThrTrpGluArgThyTyrIleuYsaPleGluValIleThrSerTrpPheGln 567
Db      1921 ---TGCACACACTGCGCTTGCTTACACTGAGAAAGTGTGTGCAGAGACCGTGTGTGTGAGAG 1978
Qy      568 SerThrVal 570
Db      1978 AGCGGCGGTG 1986

RESULT 14
US-10-152-319A-2048
: Sequence 2048, Application US/10152319A
: Publication No. US20040072160A1
: GENERAL INFORMATION:
: APPLICANT: Mendick, Donna
: APPLICANT: Porter, Mark
: APPLICANT: Johnson, Kory
: APPLICANT: Higgs, Brandon
: APPLICANT: Casale, Arthur
: APPLICANT: Elashoff, Michael
: TITLE OF INVENTION: Molecular Toxicology Modeling
: FILE REFERENCE: 44921-5089-US
: CURRENT APPLICATION NUMBER: US/10/152,319A
: PRIOR FILING DATE: 2002-05-22
: PRIOR APPLICATION NUMBER: US 60/292,335
: PRIOR FILING DATE: 2001-05-22
: PRIOR APPLICATION NUMBER: US 60/297,523
: PRIOR FILING DATE: 2001-06-13
: PRIOR APPLICATION NUMBER: US 60/298,925
: PRIOR FILING DATE: 2001-06-19
: PRIOR APPLICATION NUMBER: US 60/303,810
: PRIOR FILING DATE: 2001-07-10
: PRIOR APPLICATION NUMBER: US 60/303,807
: PRIOR FILING DATE: 2001-07-10
: PRIOR APPLICATION NUMBER: US 60/303,808
: PRIOR FILING DATE: 2001-07-10
: PRIOR APPLICATION NUMBER: US 60/315,047
: PRIOR FILING DATE: 2001-08-28
: PRIOR APPLICATION NUMBER: US 60/324,928
: PRIOR FILING DATE: 2001-09-27
: PRIOR APPLICATION NUMBER: US 60/330,867
: PRIOR FILING DATE: 2001-11-01
: PRIOR APPLICATION NUMBER: US 60/330,462
: PRIOR FILING DATE: 2001-10-22
: Remaining Prior Application data removed - See File Wrapper or PALM.
: NUMBER OF SEQ ID NOS: 2221
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 2048
: LENGTH: 4543
: TYPE: DNA
: ORGANISM: Rattus norvegicus
: FEATURE:
: OTHER INFORMATION: Genbank Accession No. NM_053527
US-10-152-319A-2048

Alignment Scores:
Pred. NO.:      3,21e-60      Length:      4543
Score:          744.00      Matches:      318
Percent Similarity: 40.1%      Conservative: 159
Best Local Similarity: 26.7%      Mismatches: 405
Query Match:      13.6%      Indels:      313
DB:              8          Gaps:          41
US-09-555-342B-2 (1-1045) x US-10-152-319A-2048 (1-4543)

Qy      6 GlnArgProThrProPheGlySerArgIleuGlyAlaProGluAuaSerGlyIleSerThrIleu 25
Db      275 CAGCGACAAAGACCGGAGCTTCGCGCCGACGCTCGAAACGCTTGAGTATCAGCACTT 334

```

QY	26	GlutrglglcInlySPProPProthnProSeGlylyLeu-----	39
		:::	
DB	335	TGAAGACGATAGCTCTCTCAGAAAGTCAATCCAGAGCAAACTTCCACGGTCCCATTTGAA	394
QY	40	-----ValSerIleLySIlleGlnMetLeuSparThrIngl	52
		:::	
DB	395	GATCGTGAAGAAAGCCTAAAGCATGACGTGCAAAAGTACGCTTCTGGATGGATCTAGTA	454
QY	52	uAlaPheGluValProGlnArgAlaProGlyLySValIleuLeuSparAlaValCySAsnHi	72
		:::	
DB	455	CGCGTGTGATGTATGAATTAAGCCCTCCCGAGGCAAGATGCTGTTTACAAAGTGTGCAACA	514
QY	72	AlauSlnLeuValGluGlyAspTryrPheGlyLeuGluPheProAspHislyrSlySileth	92
		:::	
DB	515	TTGTGAATCTGTGAGAAAAGACTACTTGGCTCACTGATGAGACGCAAGAAATCCAGA	574
QY	92	rValTrpLeuSparLeuLySProIleValLySglnIleArgTrpProLyS--HisVal	111
DB	575	GAATCGTGTGAGCCCTGCTAAGAAATTAAAGACATGTGAAAGTGAGTGGTCTGGCAT--	632
QY	111	lValVallySparheValVallySParPheProProAspHisThrGlnLeuGlnGluGlu	131
		:::	
DB	633	----TTTTCAATTATGTGAAGATTTCACCAACAGACCTCCGACGCTATCAGAAATAT	688
QY	131	uThrArgTryrLeuPheAlaLeuGlnValLySglnAspLeuAlaGlnIlyArgLeuThrCy	151
		:::	
DB	669	CACCAAGGACTACTCTGCTGCTGACAGCTCCGAGATGATTTGTGTGGAACCCCTACCTG	748
QY	151	AsnAspThrSerAlaLeuLeuLeuIleSerHisIleValGlnSerGluIleGlyAspPh	171
		:::	
DB	749	TTTCCTCGTGACCTCGCTCGCTGGGCTCTTCAACAGTGCAGAGCTGTGGGGACTA	808
QY	171	eAsp--GluAlaLeuSparArgGluHisLeuAlaLySAsnlySyrIleProGlnGln--	189
		:::	
DB	809	TGATCTCGATGAATGTGGAATGACTACATTAGCGAGTTCGCGCTTGGCCCAAAACACAC	868
QY	190	-AspAlaLeuGluAspLySileValGluPheHisHisAsnHisIleGlyGlnThrProAl	209
		:::	
DB	869	GAAGAATCTGAGAGATTAAGTATCGAGCTGCACAAAGGCCACAGAGGAATGAGCGCCAGC	928
QY	209	AGLuserAspPheGlnLeuLeuGluIleAlaArgArgLeuGlnMetTryrGlyIleArgLe	229
		:::	
DB	929	CGAGCGAGAGATGCACTTCTCGAGAAATGCCAAAAGCTCTCCATGTATGGGGTAGACTT	988
QY	229	uHisProAlaLySAspArgGluGlyThrLySIlleAsnLeuAlaValAlaAsnThrGlyI	249
		:::	
DB	989	ACACCAACGCCAAGAGATTCGAAAGCGCTGAGAAATCATGTATGGGGTCTGTGCCACGGCT	1048
QY	249	eLeuValPheGlnGlyPheThrLySIlleAsnAlaPheAsnTrpAlaLySValArgLySLe	269
		:::	
DB	1049	GTTATATCATGTGACCGGCTTGCAAATTAACAGATTTGCTTGGGCCCAAGGCTCTTAATAAT	1108
QY	269	uSerPheLySArgLySArgPheLeuIleLySLeuArgProAspAlaAsnSerAlaTryrG	289
		:::	
DB	1109	TTCTTATTAACGGAAACAATTTTACATTAAATCTCGGCGCAGAGAGTTGAACAATTGCA	1168
QY	289	nAspThrLeuGlnPheLeuMetAlaSerArgAspPheCyLeLySerPheTrpLySileCy	309
		:::	
DB	1169	AAGCAACCATTTGGCTTCAAGTTGCGCGCAACATAGAAGGCTGCCAAAGCGCTCTGGAAGATAT	1228
QY	309	eValGlnHisHisAlaPhePheArgLeuPheGlnGluProLySArgProLySProLySProVa	329
		:::	
DB	1229	TGTGGAAATATATCAATTTTTCAGACTAGT--CTACCGAAGAACCTCCAAAGAAATT	1285
QY	329	lLeuPheSerArgLySArgSerPheArgPheSerGlyArgThrGlnLySglnValLeuAs	349
		:::	
DB	1286	CCTG--ACCTGGGCTTCAAGTTCGCGTACAGTGGCAGAAACCAAGGCGCAGAGAGAG	1342
QY	349	pTryrVallySglnGlyLySlySlySValGlnPheGlnArgLySHisSerLySileHis	369
		:::	
DB	1343	AGCCAGTGCATGTGATAGACCGCGCGGCGCTTACTTGGAAAGCTCTGTCACAGCAACGATA	1402

```

Qy 369 sserile---ArgSerLeuAlaSerGlnProThrGluLeuAsnSerGluValLeuGluG1 388
Db 1403 CACCATCTCTCGCAGCTTGTGATGAGCCTCAGTAAATGAAAACATGAAATATATC----- 1457
Qy 388 nserGlnInserTherSerLeuThrPheGlyGluGlyAlaGluSerProGlyGlyInse 408
Db 1457 ----- 1457
Qy 408 rcyArxArgGlyLygGlyProLybValSerAlaGlyGluProGlySer-HisProSerP 428
Db 1458 -----ATGAAGATTCTATGTCTGTCTGCAGAGTTGTACAGGCCAGTAAGC 1504
Qy 428 roAlaProArxArgSerProAlaGlyAsnLybGlnAlaAspGlyAlaAlaSerAlaProT 448
Db 1505 CACACCAAGGTCTATCTCT-CAGACCAACGTGATCAC-----ACTGTACCCCGAGA 1557
Qy 448 hrtGluGluGluGluValAlaLybAsp-ArgThrGlnInserLybProGlnProPro 467
Db 1558 AAAAGGCTGAGAGACACGTCGAGAGAGAGACAGAGAGAGAGAGAGAGAGAGAGAGAG 1617
Qy 468 GlbProSerThrGlySerLeuThrGlySerProHisLeuSerGluLeuSerValAsnSer 487
Db 1618 CGCCCGTCGAGACCCCTCGGCGACAGGAGAAATCAC----- 1653
Qy 488 GlnGlyLybAlaProAlaAsnValThrLeuSerPro-----AsnLeuSerPro 504
Db 1654 ---CAGGGCATGCTCTGACATCATGTCCC-CTGTCACTTCATCAGGCCCATCTTGACCT 1709
Qy 505 -----AspThrLybGlnAlaSerProLeuIle 513
Db 1710 CCTCTCCACAGAGCTCCGTAGAGAGTGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1763
Qy 514 SerProLeuLeuAsnAspGlnAlaCybProArxThrAspAspGluAspGluGlyArgArg 533
Db 1764 -----GGTTCTGCTGTGCTTACAGCAGAG----- 1787
Qy 534 LybArxPheProThrAspLybAlaTybPheIleAlaLybGluValSerThrThrGluArg 553
Db 1788 -----CTGTCTCGAGAGCTGATGCCCAGAGGAG 1817
Qy 554 ThrTybLeuLybAspLeuGluValIleThrSerThrPheGln----- 567
Db 1818 GCCCTACCTAGGGGACCAAGAGTGGCATTTAGTACAGGCGCGGTGGCAAGAAC 1877
Qy 568 SerThrAlaSerLybGluAspAlaMetProGluAlaLeuLybSerLeuIle----- 584
Db 1878 ACACTGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1937
Qy 585 -----PheProAsnPheGluProLeuHisLybPheHisThrAsnPheLeuLyb 600
Db 1938 GGGTACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1982
Qy 601 GluIleGluGlnAlaLeuThrGluGlyArgSerAsnAlaGlnIleArgAspTyr 620
Db 1983 CAG----- 1985
Qy 621 GlnArgIleGlyAspValMetLeuLybAsnIleGlnGlyMetLybHisLeuAlaAlaHis 640
Db 1985 ----- 1985
Qy 641 LeuTrpLybHisSerGluAlaLeuGluAlaLeuGluAlaLeuLybSerSerArgArg 660
Db 1986 AGTTGGCAGAGCTTCTTGCC-----ATCCGCTCGCCCTCTCTCTCTCTCTCTCTCTCT 2024
Qy 661 LeuGluAsnPheCybArxAspPheGluLeuGlnLybValCybTybTrpLeuProLeuAsnThr 680
Db 2025 CTTCCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2080
Qy 681 PheLeuLeuArgProLeuHisArgLeuMetHisTybLybGlnVal----- 695
Db 2081 -----ATAGCCCT---CACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2131
Qy 696 -----LeuGluArgLeuCybLyb-HisHisProProSerHisAlaAspPh 710

```

```

Db 2132 GCCCAAGCGGCTCTCTTGAGCGCTCCCTAGACATGACCCGAGTGACAGTTCCAGAG-- 2189
Qy 710 eArxAspCybArxAlaAlaLeuAlaGluIleThrGluMetValAla----- 725
Db 2190 -----GAAAGACTGACGTAGCGAGACAGACTGCGGCTGATGAGAGAGACAG 2239
Qy 726 -----GlnLeuHisGlyThrMetCileYMetGluAsnPheGlnLybLe 740
Db 2240 TGCCACAGACTCGGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2299
Qy 740 uHis---GluLeuLybLybAspLeuIleGlyIleAspAsnLeuValProGlyArgG1 759
Db 2300 TCAAGATGAGCTGTGTAACACCAACCAATATATACGAATG-----AAAAGAAC 2350
Qy 759 uPheIleArgLeuGlySerLeuSerLybLeuSerGlyLybGlyLeuGlnGlnArgMetPh 779
Db 2351 CTCTTTAGAAACCTTACAGAAACCTGCTTACAAACGAATGGAGAAAGAG----- 2402
Qy 779 ePheLeuPheAsnAspValLeuLeuTybThrSerArgGlyLeuThrAlaSerAsnGlnPh 799
Db 2403 -----CTTCCACATCCCGTGGCTGGCTGCTCCAGGACAGA 2440
Qy 799 eLybValHisGlyGlnLeuProLeuTybGlyMetThrIleGluGluSerGluAspGluTr 819
Db 2441 GATGCGGCCCATGATTGAGCACTGTACCTGAGAGAGATTAACAAATCTTCTGCGAGAA 2500
Qy 819 P-----GlyValProHisCybLeuThrLeuArgGlyGlnArgGlnSerIleIleVa 836
Db 2501 GCTCATGATAGTGTCCGAGATCTCTCAGCTATATAGACTGTGCGAGAGAACCCACAGATT 2560
Qy 836 lAlaAlaSerSerArgSerGluMetGluTybTrpValGluAspIleGlnMetAlaIleAs 856
Db 2561 CATAGAGAGGGTTTCACTACTACTACCAAGAGCTGGTGGCAGAAATCGAAAGAAAGACTGA 2620
Qy 856 pLeuAlaGluLybSerSerSerProAlaPro----- 866
Db 2621 GACCATTTGAAACAGAGATGAGAACCCACCCACCTCAGCCCTCAGCAGAGAGAGGT 2680
Qy 866 ----- 866
Db 2681 GTTTCAGAGAAACGCTGCTGTAGAGAGAGACGATGATGACGTGACGCCAGTGGGA 2740
Qy 867 -----GluPheLeuAlaSerSerProProAsnLy 877
Db 2741 CGCTTCTCACACAGCCAGGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 2800
Qy 877 sSerProAsp----- 880
Db 2801 CAGTGGGAATGGAGAGAGGCTTTCTGTGACAGAGACAGCTMAAGAACAGAGAGAGA 2860
Qy 881 -----GluAlaThrAlaAlaAspGlnIle-----SerGluAs 891
Db 2861 GAGAGCTGACAGATCTGCCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2920
Qy 891 pAspLeuSerAlaSerArgThrSer-----LeuGluArgGlnAlaProHisArgGlyAs 909
Db 2921 TCAAGTGTGGCCATGCATCTCTCTGAGGTTTGAACAAAGTCT----- 2966
Qy 909 nThrMetValHisValCybTrpHisArgAsnThrSerValSerMetValAspPheSerI1 929
Db 2967 -----CATTTGAGTGTGCTCAGCTGTAAGGTAGAA---ACCAT 3001
Qy 929 eAlaValGluAsnGlnLeuSerGlyAsnLeuLeuArgLybPheLybAsnSerAsnGlyTr 949
Db 3002 AAGTGTGGCAGAGTCTCTCCAGAGAGAGTg-----AAGCTAGACCTTAC 3049
Qy 949 pGlnLybLeuTrpValAlaPheThrAsnPheCybLeuPhePheTybSerHisIleGlnAs 969
Db 3050 CAGAGAGGTGCCAGTGTCCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3109
Qy 969 pAsnHisProLeuAlaSerLeuPro-----LeuLeuGlyTybSerLeuThrIleProSe 987

```

Dh 3110 CGAT---CCTGGGGAGATCTGGAGCCAGGCTGTCTATAGACGCCCAAAGATCAGTC 3166
Qy 987 rglserglunilleglnlyshptryvalphelyserhlsvalty 1007
Dh 3167 TGAACCCAGGAGTACTACACACCACCAATATACCAAACTGAAAGAGGATTC 3226
Qy 1007 rtyrpherglaglu-serglu-----tyrtrpherg 1018
Dh 3227 GGAAGAGGAGTGTGAAGGAGATGTGATCATCGGAGATGCCGATTCATGACCA 3286
Qy 1018 luatrgpmetgluvalle-----ArgSerAlaThrSerSerAlaSerArgProH 1035
Dh 3287 GGGCTGGCTCAGGCAATTAAGAGCCCAAGAGAGACACCCGACATGTCAATACCA 3346
Qy 1035 tsvalleuserhlslyserglu 1042
Dh 3347 A-GTAGGTGTCCATTAAGAGACA 3368
RESULT 15
US-11-036-196-2048
Sequence 2048, Application US/11036196
GENERAL INFORMATION:
APPLICANT: Mendrick, Donna
APPLICANT: Johnson, Kory
APPLICANT: Higgs, Brandon
APPLICANT: Casale, Arthur
APPLICANT: Elashoff, Michael
TITLE OF INVENTION: Molecular Toxicology Modeling
FILE REFERENCE: 44921-5089-US
CURRENT APPLICATION NUMBER: US/11/036,196
CURRENT FILING DATE: 2005-01-18
PRIOR APPLICATION NUMBER: US/10/152,319
PRIOR FILING DATE: 2002-05-22
PRIOR APPLICATION NUMBER: US 60/292,335
PRIOR FILING DATE: 2001-05-22
PRIOR APPLICATION NUMBER: US 60/297,523
PRIOR FILING DATE: 2001-06-13
PRIOR APPLICATION NUMBER: US 60/298,925
PRIOR FILING DATE: 2001-06-19
PRIOR APPLICATION NUMBER: US 60/303,810
PRIOR FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: US 60/303,807
PRIOR FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: US 60/303,808
PRIOR FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: US 60/315,047
PRIOR FILING DATE: 2001-08-28
PRIOR APPLICATION NUMBER: US 60/324,928
PRIOR FILING DATE: 2001-09-27
PRIOR APPLICATION NUMBER: US 60/330,867
PRIOR FILING DATE: 2001-11-01
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 2221
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2048
LENGTH: 4543
TYPE: DNA
ORGANISM: Rattus norvegicus
FEATURE:
OTHER INFORMATION: Genbank Accession No. NM_053927
US-11-036-196-2048
Alignment Scores:
Pred. No.: 3,21e-60 Length: 4543
Score: 744.00 Matches: 318
Percent Similarity: 40.1% Conservative: 159
Best Local Similarity: 26.7% Mismatches: 405
Query Match: 13.6% Indels: 313
Gaps: 41
US-09-555-342B-2 (1-1045) X US-11-036-196-2048 (1-4543)

Qy 6 GluArgProThrProGlySerArgLeuGlyAlaProGluAsnSerGlyIleSerThrLeu 25
Dh 275 CAGGACAAAGACCGGGACCTTCGCGCGGCGGACGCAAAACGCTGATGATCAGCAGT 334
Qy 26 -GluArgGlyGlnIlySerProProProThrProSerGlyIlyLeu----- 39
Dh 335 TGAAGAGATAGTCTCTCAGAACTCATCCAGAGCAAACTCTCAGGCTCCCATGGA 394
Qy 40 -----ValSerIleIleGlnMetLeuAspThrGlnI 52
Dh 395 GATCGTAAAAAGGCTAAAGACATGACGTAAGATGACGCTTCGATGATGATCGATGA 454
Qy 52 ualaphelgluvalproGlnArgAlaProGlyIlyValLeuAspAlaValCysAsnH 72
Dh 455 CGGCTGTGATGTATAGGCGCTCCAGAGCAAGCTGTGTAAACAAAGTTCGAAACA 514
Qy 72 sLeuAsnLeuValGluGlyAspTyrPheGlyLeuGluPheProAspHisIlyIleTh 92
Dh 515 TCTGAACCTGTGAAAAAGACTACTTCGCTCTCAGATGAGACCGACAGATTCGAA 574
Qy 92 rValTrpLeuAspLeuLeuIysProIleValIyGlnIleArgArgProIys--HisVa 111
Dh 575 GAACGTGGTGGACCTGCTAAGAAATTAAACAGATTGCAAGTGGTGGCAT-- 632
Qy 111 lValIlyAspPheValIlyIysPheProProAspHisThrGlnLeuGlnIlyGlu 131
Dh 633 ----TTTTCATTATATGTGAAGTTTAACTCCACAGACCTGCGCATTCAGAAATAT 688
Qy 131 uThrArgTyrLeuPheAlaLeuGlnValIyGlnAspLeuAlaGlnIlyArgLeuThrCy 151
Dh 689 CACGAGTACTACTCTGCTTGCAGCTGCCAGATGATATGTGCTGAGCCCTACCTG 748
Qy 151 sAsnAspThrSerAlaIleLeuLeuIleSerHisIleValIyGlnSerGluIleGlyAsp 171
Dh 749 TTCCTGTGACCTCGCCCTGCTGCGCTCTCAACAGTCAGACAGCTTGGGGACTA 808
Qy 171 eAsp---GluAlaLeuAspArgGluHisIleAlaIlyAsnIlyIleProGlnGln-- 189
Dh 809 TGATCTGTAGTAAATGTGAATGCAATGCAATTGAGAGTTCGCTTGCAGCAACAC 868
Qy 190 -AspAlaLeuGluIlyAspIlyIleValIyGlnPheHisIleAsnHisIleGlyGlnThrProAl 209
Dh 869 GAAAGAACTGGAGATTAAGTATGATCGAGCTGCACAAAGCCACAGAGAAATGACCGAC 928
Qy 209 agIySerAspPheGlnLeuLeuGluIlyAlaArgArgLeuGluMetIyGlyIleArg 229
Dh 929 CGAGCAGAGATGCACTTCCTGGAGATGCAAAAGCTCCATGTATGGGTAGACTT 988
Qy 229 uHisProAlaIlyAspArgGluGlyIlyIlyHisIleAsnLeuAlaValAlaAsnThrGlyI 249
Dh 989 ACACCAAGCAAAAGATTCGAAAGCGTGAGATCATGTGTGGGTCTGTGCCACGGCT 1048
Qy 249 eLeuValPheGlnIlyPheThrIlyIleAsnAlaPheAsnTrpAlaIysValArgIlySle 269
Dh 1049 GTTATCTATCGTACCGGCTTCGAAATTAACAGATTGCTTGGCCCAAGTCTTAATAAT 1108
Qy 269 uSerPheIlyAspGlyAspPheLeuIlyIleValIyGlnArgProAspAlaAsnSerAlaIyrgI 289
Dh 1109 TTTCTATTAACGGAACAACTTTTACATTAATATCCGCGCAGAGAGTTGAAACAATTGCA 1168
Qy 289 nAspThrLeuGluPheLeuMetAlaSerArgAspPheCylIySerPheThrIlyIlyIleCy 309
Dh 1169 AAGCACCATTGGCTTCAAGTTGCCGAACATAGGCTGCCAAGGCGCTCGGAAAGATAG 1228
Qy 309 sValGluHisIleAlaPhePheArgLeuPheGluGluProIlyProIlySerProIysProVa 329
Dh 1229 TGTGAACATCATCATTTTTCAGACTACG---CTACAGAAAGCACTCCAAAGAAATT 1285
Qy 329 lLeuPheSerArgIlySerSerPheArgPheSerGlyIlyArgThrGlnIlyGlnValLeuAs 349
Dh 1286 CTG---ACCTGGGCTCAAGTTCCGCTACAGTGGCAGAAACCGACGCGACGAGAG 1342

QY	349	pTYrVallysguLgylgYhlsElyblyValGlnpHeGluArgylhlsErylIlehi	369
Db	1343	AGCCAGTCATTTGATGATGAGCCGCCGCCGCTTATTCGAAAGCGTCGTGCACCAACGATA	14020
QY	369	eserIle---ArgserLeuAlaSerGlnProthrGluLeuAmsSerGluAlaLeuGluG	388
Db	1403	CACCATGTCTGCGACCTTGATGAGGCTCAGTGAATGAAAACATGAAATATAC----	1457
QY	388	nSerGlnGlnSerThrSerLeuThrPheGlyGluGlyAlaGluSerProGlyGlyGlnSe	408
Db	1457	-----	1457
QY	408	rCyArGArgGlyblysgLuproLyValSerAlaGlyLupProGlySer-HisProSerP	428
Db	1458	-----ATGAAGATTCTATCTGTGCTGCACAGGTGGTATCAGGCCAGCTATGC	15040
QY	428	roAlaProArGArgSerProAlaGlyAmsLyGlnAlaAspGlyAlaAlaSerAlaProT	448
Db	1505	CACAAACCAAGGTCACTCTC-CAGACCAACCTGATACC-----ACTGTGACCCCGGAGA	15570
QY	448	hrgLugLugLugLugLuvAlVallyAsp-ArgThrGlnGlnSerLyProGlnProPro	467
Db	1558	AAAGAGCTGAGGAGGAGCAGCGGAGGAAGAGACAGAAAGAAAGAGCCAGAGACCA	16170
QY	468	GlnProSerThrGlySerLeuThrGlySerProHisLeuSerGluLeuSerValAmsSer	487
Db	1618	CGCCCGGTGCGAGCCCTCCGGCGACGAGGAAAGTCCAC-----	16530
QY	488	GlnGlyGlyValAlaProAlaAmsValThrLeuSerPro-----AsnLeuSerPro	504
Db	1654	---CAGGCGATGCGCTCTGACTATGTACC-CGTGCACTCCATCAGCCCATCTTGACCT	17090
QY	505	-----AspThrLySGlnAlaSerProLeuIle	513
Db	1710	CCCTCTCCACAGACCTCCGTAGAGGCGCAAGAGAAAGAGGCCCTTCCA-----	17630
QY	514	SerProLeuLeuAmsAspGlnAlaCySProArGThrAspAspGluAspGlyAlaArgArG	533
Db	1764	-----GGTTCTGCTGTGCTTGACGACGAG-----	17870
QY	534	LySArGpHeProThrAspLyAlaIaTyrrHelleAlaLySGluValSerThrThGluArg	553
Db	1788	-----CTGTCTCGAGGTGTGATGCCAGGGGAG	18170
QY	554	ThrTyrrLeuLyAspLeuGluValIleThrSerTrpPheGln-----	567
Db	1818	GCCTACCTTAGGGACCAAGATGTGGCATTTAGTTACAGCAGCCGGCTGGCAAGGAAC	18770
QY	568	SerThrValSerLySGluAspAlaMetProGluAlaLeuLySerLeuIle-----	584
Db	1878	ACACTGTTTTCTCTCTCCGTGACCTCCAGAGTCAATCCCTCGCTCTTAGTAAGAC	19370
QY	585	-----PheProAmsPheGluProLeuHisLySPheHisThrAspPheLeuLyS	600
Db	1938	GGGTACCTCTCTTTCCCAACCTGTCT-----GAAACCAACCTCTGCCC	19820
QY	601	GluIleGluGlnArgLeuAlaLeuTrpGluGlyArgSerAmsAlaGlnIleArgAspTrx	620
Db	1983	CAG-----	1985
QY	621	GlnArgIleGlyAspValMetLeuLyAsnIleGlnGlyMetLyblyAlaAlaHis	640
Db	1985	-----	1985
QY	641	LeuTrpLyblySerGluAlaLeuGluAlaLeuGluAmsGlyIleLybSerSerArgArg	660
Db	1986	AGTTGGAGCAGATTCTTGCCC-----ATCCGCTGCCCTCCCTC	20240
QY	661	LeuGluAmsPheCyArgAspPheGluLeuGlnLyValCyTyrrLeuProLeuAmsThr	680
Db	2025	CTTCCCTGCTCTTCTTATCTTCTT-CTTCCGTGCTGTGCTCTT---CTCAGTACC	20800
QY	681	PheLeuLeuArgProLeuHisArgLeuMetHisTyrrLybGlnVal-----	695

Db	2081	-----ATAGCCCT-----CACTCTCTCTCCCTCTGAGCTCTGCTCTGCTCTCTGGA	2131
Qy	696	-----LeuGluArgLeuCysIys-HisH1sProProSerHisValAspPh	710
Db	2132	GCCCAAGCGCGCTCTTGAAGCGCTCCCTAGACAATGACCCGAGTGCAGTTCAAG--	2189
Qy	710	eArgAspCysArgValAlaLeuAlaGluIleThrGluMetValAla-----	725
Db	2190	-----GAAAGACTGACAGTGTGACGGACAGACATCGCGCTTAGATGAGAGACAG	2239
Qy	726	-----GluLeuH1sGlyThrMetIleGlyMetGluAsnPhenGlyIle	740
Db	2240	TGCCACAGAGTCCGACCAAGAGGAGAGACGACAGCTCAAGGCACAGAGCTAGATTAAC	2299
Qy	740	uHis-----GluLeuIlyIysAspLeuIleGlyIleAspAsnLeuValAlaProGlyArgI	759
Db	2300	TCAAGATGAGTGTGTAACACCAACCAATATTAGGAAGCTG-----AAAGAAC	2350
Qy	759	uPheIleArgLeuGlySerLeuSerIlyLeuSerGlyIlyGlyLeuGlnGlnArgMetPh	779
Db	2351	CTTCTTGAACACTCTCAAGAAATGCGCTTAAACAAGAAAGGAGGAGGAGG-----	2402
Qy	779	ePheLeuPheAsnAspValLeuLeuIlyThrSerArgGlyLeuThrAlaSerAsnGlnPh	799
Db	2403	-----CTCTCACATCCCGTCGCGCTGCTGCTGCAGCAGCA	2440
Qy	799	eIysValH1sGlyIleuProLeuIlyGlyMetThrIleGluGluSerGluAspGluTr	819
Db	2441	GGATGCCGCCCATGATTTAGCCACTTGTACCTGAGAGAACCTAAACAATCTTGGGAGAA	2500
Qy	819	p-----GlyValProHisCysLeuThrLeuArgGlyGlnArgIleSerIleIleVal	836
Db	2501	GCTCATGATGAGTCCGAGAGATCCCTCAGTATATGAGTGCAGAGAAACCAAGAGTT	2560
Qy	836	lAlaAlaSerSerArgSerGluMetGluIlyIleTrpValGluAspIleGluMetAlaIleAs	856
Db	2561	CATAGGAGGGGTTTCATCTACTACCCAGACAGCTGGGTCAAAACTGGAAACGAAGACTGA	2620
Qy	856	pLeuAlaGlyIlySerSerSerProAlaPro-----	866
Db	2621	GACCATTTGAACAGAGGTGAACCCACCCACACCTCAGCCCTCAGCAGGAAAGGT	2680
Qy	866	-----	866
Db	2681	GTTGCAGAGAAACGCTGCTGTGTAGAGAGAGACAGTATGAACGTGACCGCAGTGGGA	2740
Qy	867	-----GluPheLeuAlaSerSerProProAsnAsnIly	877
Db	2741	CGCTTTCACACAGCCAGGAGTGTGTGATGACCAAGAAATCAGACCCCGACAGCCGTCA	2800
Qy	877	sSerProAsp-----	880
Db	2801	CAGTGGGAATGGAGAGGCGCTTTTGTGACAGAAACAGCTAAGAAACAGAGAGAGA	2860
Qy	881	-----GluAlaThrAlaIleAspGlnIly-----SerGluAs	891
Db	2861	GGAAGCTGACAGAGCTGCGCCCGGAGCAGAAACAACAGCACTGTCTCCACAGAGAGAGA	2920
Qy	891	pAspLeuSerAlaSerArgThrSer-----LeuGluArgGlnAlaProHisArgGlyAs	909
Db	2921	TCAAGTGTCCGCCATTCACATCTTCTGTGAGGTTTGGAAACAAGACT-----	2966
Qy	909	nThrMetValHisValCysThrPhsArgAsnThrSerValSerMetValAspPheSerI	929
Db	2967	-----CATTTTGAAGTCCCTCAGCTGTGAAGGTAGAA-----ACCAT	3001
Qy	929	eAlaValGluAsnGlnLeuSerGlyAsnLeuLeuArgIlyPheIlyAsnSerAsnGlyTr	949
Db	3002	AAGCTGTGGCAGTGTCTCTCCAGAGAGAGT-----AAGCTAAGACCTCTAC	3049
Qy	949	pGlnIlyLeuTrpValValPheThrAsnPheCysLeuPhePheTrpIlySerHisGlnAs	969

```
Db      3050 CAAGAGGTGCCAGTGTCCACACAGAGACGAAACCATCATCGTACGAGCTTCTCAGGT 3109
QY      969 pAenHiSProleuLaseSerLeuPro-----LeuLeuGlyTyrSerLeuThrIleProSe 987
Db      3110 CGAT--CTGGGGCAGATCTGGAGCCAGGTGTGCTTAATGAGCGCCCAACGATCACGTC 3166
QY      987 rGluSerGluAsnIleGlnLysAerTyrValPheLysLeuHisPheLysSerHisValTy 1007
Db      3167 TGAATCCACGAGTACTACACACACATATCATCCAAACTGTGAAGAGAGGATTC 3226
QY      1007 rTyrPheArgAlaGlu-SerGlu-----TyrThrPheG 1018
Db      3227 GGAGACGAGGATTGAGAGCGAATAGTCATCAGGAGATGCCGACATTGACCATGACCA 3286
QY      1018 lArgTrpMetGluValIle-----ArgSerAlaThrSerSerAlaSerArgProH 1035
Db      3287 GCGCGTGGCTCAGGCAATTAAAGAGGCCAAAGAGCAGACCCCGACATGTCAGTAAACCA 3346
QY      1035 lsvValLeuSerHisLysGluSer 1042
Db      3347 A-GTAGTGTGTCATTAAGAGACA 3368
```

Search completed: May 20, 2006, 04:14:36
Job time : 2901 secs

This Page Blank (uspto)

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - nucleic search, using frame_p2n model

Run on: May 20, 2006, 03:27:59 ; Search time 30 Seconds
(without alignments)
3686.546 Million cell updates/sec

Title: US-09-555-342B-2

Perfect score: 5463
Sequence: 1 MGELEQRPTPSRLGAPENS.....SATSSASRPHVLSKESLVY 1045

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 224314 seqs, 35277956 residues

Total number of hits satisfying chosen parameters: 448628

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame+g2n.model -DEV=xlh
-Q=/abs/ABSSWB.epool/US09555342/runat.19052006.121512.13499/app.query.fasta.1
-DB=Published Applications NA New -QFMT=faastap -SUFFIX=p2n.rnpbn -MINMATCH=0.1
-LOOPCL=0 -LOGEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=01osum62
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -HOST=abes02h
-USER=US09555342.ecgn.1.1.13 @runat.19052006.121512.13499 -NCPU=6 -ICPU=3
-NO MAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THRADs=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications_NA_New:*
1: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09_NEW_PUB_seq.*
2: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US06_NEW_PUB_seq.*
3: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US07_NEW_PUB_seq.*
4: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US08_NEW_PUB_seq.*
5: /EMC_Celerra_SIDS3/ptodata/2/pubpna/PCT_NEW_PUB_seq.*
6: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10_NEW_PUB_seq.*
7: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11_NEW_PUB_seq.*
8: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US60_NEW_PUB_seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	713	13.1	2867	6	US-10-511-937-344 Sequence 344, App
2	396.5	7.3	3953	6	US-10-473-173-56 Sequence 56, Appl
3	176.5	3.2	6469	7	US-10-505-928-228 Sequence 228, App
4	173	3.2	2559	7	US-11-217-529-2747 Sequence 2747, App
5	169	3.1	646	6	US-10-488-618-2736 Sequence 2736, App
6	155.5	2.8	4119	7	US-11-217-529-77046 Sequence 77046, A
7	155	2.8	7430	6	US-10-559-415-1 Sequence 1, Appl
8	152	2.8	7465	6	US-10-505-928-311 Sequence 311, App
9	151.5	2.8	3054	7	US-11-217-529-5019 Sequence 5019, App

10	150	2.7	6505	6	US-10-505-928-721	Sequence 721, App
11	140	2.6	7268	6	US-10-559-415-189	Sequence 189, App
12	137.5	2.5	11185	6	US-10-505-928-450	Sequence 450, App
13	137	2.5	11128	7	US-11-118-524-1	Sequence 1, Appl
14	136.5	2.5	5040	7	US-11-217-529-77055	Sequence 77055, A
15	136	2.5	1617	7	US-11-217-529-173406	Sequence 173406, A
16	133.5	2.4	1863	7	US-11-217-529-75500	Sequence 75500, A
17	133.5	2.4	1866	7	US-11-217-529-148	Sequence 148, App
18	133.5	2.4	4080	7	US-11-217-529-1845	Sequence 1845, App
19	133	2.4	3210	7	US-11-217-529-80763	Sequence 80763, A
20	131.5	2.4	4263	7	US-11-327-900-5	Sequence 5, Appl
21	131.5	2.4	5449	6	US-10-505-928-800	Sequence 800, App
22	128	2.3	13515	6	US-10-505-928-358	Sequence 358, App
23	127	2.3	3246	7	US-11-217-529-78477	Sequence 78477, A
24	127	2.3	3318	7	US-11-217-529-4920	Sequence 4920, App
25	127	2.3	3444	7	US-11-217-529-5473	Sequence 5473, App
26	127	2.3	3658	7	US-11-217-529-6026	Sequence 6026, App
27	126.5	2.3	3210	7	US-11-217-529-2695	Sequence 2695, App
28	126	2.3	10211	6	US-10-505-928-326	Sequence 326, App
29	125	2.3	3705	6	US-10-505-928-655	Sequence 655, App
30	124.5	2.3	3854	6	US-10-511-937-549	Sequence 549, App
31	122.5	2.2	3261	7	US-11-217-529-1461	Sequence 1461, App
32	122	2.2	2358	7	US-11-217-529-190978	Sequence 190978, A
33	122	2.2	3387	7	US-11-217-529-78908	Sequence 78908, A
34	121.5	2.2	3117	7	US-11-217-529-774	Sequence 774, App
35	121.5	2.2	3729	7	US-11-217-529-79062	Sequence 79062, A
36	120.5	2.2	4053	7	US-11-217-529-53	Sequence 53, Appl
37	120	2.2	4584	7	US-11-217-529-3073	Sequence 3073, App
38	119.5	2.2	9588	6	US-10-505-928-101	Sequence 101, App
39	119	2.2	2667	7	US-11-217-529-2499	Sequence 2499, App
40	119	2.2	3243	7	US-11-217-529-76988	Sequence 76988, A
41	118.5	2.2	3366	7	US-11-217-529-76415	Sequence 76415, A
42	118.5	2.2	3446	6	US-10-505-928-176	Sequence 176, App
43	118.5	2.2	7280	6	US-10-505-928-842	Sequence 842, App
44	118	2.2	2511	7	US-11-217-529-77149	Sequence 77149, App
45	117.5	2.2	3471	7	US-11-217-529-3423	Sequence 3423, App

ALIGNMENTS

RESULT 1
US-10-511-937-344
; Sequence 344, Application US/10511937
; Publication No. US2006008836A1
GENERAL INFORMATION:
; APPLICANT: EXPRESSION DIAGNOSTICS, INC.
; APPLICANT: Wohlgemuth, Jay
; APPLICANT: Fry, Kirk
; APPLICANT: Woodward, Robert
; APPLICANT: Ly, Ngoc
; APPLICANT: Prentice, James
; APPLICANT: Morris, Macdonald
; APPLICANT: Rosenbery, Steven
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
; TITLE OF INVENTION: AND MONITORING TRANSPLANT REJECTION
; FILE REFERENCE: 50661200104
; CURRENT APPLICATION NUMBER: US/10/511,937
; CURRENT FILING DATE: 2004-10-19
; PRIOR APPLICATION NUMBER: PCT/US2003/012946
; PRIOR FILING DATE: 2003-04-24
; PRIOR APPLICATION NUMBER: US 10/131,831
; PRIOR FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: US 10/325,899
; NUMBER OF SEQ ID NOS: 3117
; SOFTWARE: PatentIn version 3.2
; LENGTH: 2867
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-511-937-344
Alignment Scores:

Pred. No.:	1,3e-40	Length:	286
Score:	713.00	Matches:	90
Percent Similarity:	46.1%	Conservative:	90
Best Local Similarity:	31.3%	Mismatches:	197
Query Match:	13.1%	Indels:	130
DB:	6	Gaps:	17

```

; NUMBER OF SEQ ID NOS: 498
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 56
; LENGTH: 3953
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-473-173-56

Alignment Scores:
Pred. No.: 1,41e-18 Length: 3953
Score: 396.50 Matches: 146
Percent Similarity: 38.44 Conserved: 98
Best Local Similarity: 23.04 Mismatches: 245
Query Match: 7.34 Indels: 147
DB: Gaps: 20

US-09-555-342b-2 (1-1045) x US-10-473-173-56 (1-3953)
Qy 29 GlnlySPROProthProSerGly---Lysleu-----Val 40
Dy 197 AAGAAGAAGATGCCATGCCATTGGGTTGAATGAACGACCCGGCGCTACAGGTG 256
Qy 41 Serlleys-----lleglmetleuSPArpThrglnGlnlaPhe 54
Dy 257 TCACAGAACAGGTGCTGCTGTGCCGAGTCCCACTGCTTAATAAGAGATTGTGGAGTTC 316
Qy 55 GluValProGlnAglAProGlyLysValleuLeuAPRAlaValCyAsnHileuAsn 74
Dy 317 ACCCTGCGCTGGAGACAGCTGGCCGAGAAAGCCTCGAGCGCGTGGCCAGAGCTGGAG 376
Qy 75 LeuValGlnGlyAapTyPheGlyLeuGlnPheProAPRhlblyValleThrValTyr 94
Dy 377 CTGGGAGAGCTACTTACTTACGCTCTGTGACTACAAACAGCAAAATCAGCGCGGTGG 436
Qy 95 LeuAPRleuLeuLysProileValleGlnle---ArgArpProLysHileValVal 113
Dy 437 GTAGATTGGAAAAACCTTTGAAGAGAGCTGGATTAATATGACTTGAACCTACCGTTC 496
Qy 114 LysPheValValLysPhePheProAPRhlbthrglnleGlnGlnGlnLeuThrg 133
Dy 497 TATTTTGAAGTGGTGTATATGTGCTTCACTTCTCAGCTGACAGAGATTAACAG 556
Qy 134 TytleuPheAlaGlnValleGlnAPRleuLaglnGlyArgLeuThrCyAsnAPR 153
Dy 557 TATCAGATTATCTGCACACTGAAGAAAGATATCTTGAAGAGATATTCCTTGATCTTA 616
Qy 154 ThSerAlaAlaLeuLeuLysSerHileValGlnSerGlnleGlyAspPheAPR 173
Dy 617 GAACAGCAATTCAGCTAGCAGAGCTTACGCTTCAAGCGGATTTTGGATTCAG 676
Qy 174 AlaLeuAPRatGlnHileuAlaLysAPRlyTyrllePro-----Gln 188
Dy 677 TATGAATCCCGAGACTTCTTCAAGAAATTCCTTGTTCCTGGGAGTTCACAGAT 736
Qy 189 GlnAPRAlaLeuGlnAPR-----LyslleValGlnPheHlbnHleAsnHleGly 205
Dy 737 GAAGAAAGATATTGAAGAGCAACCCAAAGAGGCTTACTACATCAGAAATACAGAGG 796
Qy 206 GlnThrProLaglnSerAPRPhelGlnleuGlnleAlaArgArgLeuGlnleTyr 225
Dy 797 CTCACAGCTCTGATGTGAATGCTGATGACGAGAGTGAAGAGATGATGAGGCTAT 856
Qy 226 GlnlyleArgleuHleProAlaLysAPRatGlnGlyHlyHlylleAsnleAlaVal 245
Dy 857 GGAAGAGAGAGCTACCTGCTTAAGATGACCAAGAGATGACATATCATTCATTGAGAGCTGT 916
Qy 246 AaThrglylleleuVal---PheGlnGlyPheThrylleAsnAlaPheAsnTyrAla 264
Dy 917 CTGAAGATATCTTGTGAACACAGAAATGAAGAGCTCTGTGGATTTTAAAGTGGCAT 976
Qy 265 LysValArgLysLeuSerPheLysArgLysArgPheleuLysLysLeuArgProAPRAla 284
Dy 977 GACATTCGCAACATGTCCCAACAAAGTCTTTTTCATTAAGAGCTG-----GCA 1027

```

```

Qy 285 AaSerAlaTyrglnAPRThrleuGlnPheleuMetAlaSerArgAPRPhCyAlaLysSer 304
Dy 1028 AATAAA-----GAGAGAGCATTTCAATTAACTGAAGCATGAAACGAAATATAC 1081
Qy 305 PheTyrLyslleCyValGlnHleAlaPhePheArgleuPheGlu----- 320
Dy 1082 ATTTGAGAGCTCTGTGTGGCGACACAAAGTTTACAGACTTAACAGTGTAACTGCA 1141
Qy 321 -----GlnProLysPro 324
Dy 1142 ACTCAGACTGTCAAGTGAACCCAAATGAGAGAGAGTCTTCTTCAAGAGATCTCTGCT 1201
Qy 325 LysProLysProValleuPheSerArgGlySerSerPheArgPheSerGlyArg----- 342
Dy 1202 AAACCCAGCCCTAGTGATGCTCTCCCAACGAGATTCGACATATATGACATTATACA 1261
Qy 343 -----ThrglnLysGlnValleuAPRlyTyrllyGlnGlyLys----- 355
Dy 1262 GAACCATATGCTTCTTCCCAAGATTAACCTTTTGTGCCCAACGAAACGATACACTGT 1321
Qy 356 HileLysValGlnPheGlnArg-----LysHileSer 366
Dy 1322 CACTCTCAGACAAAGCTTGAATAGAGCCAGATTAAGCTCAACGCTCGGATCCGATATGTC 1381
Qy 367 LyslleHileSerleArgSer----- 373
Dy 1382 AGTGTCTACAGTGCACACAGACCAACTCTTAATAATATCTCAGCCCTACTTCAAGCCC 1441
Qy 374 -----LeuAlaSerGlnProThrglnleuAsnSerGlnValleu----- 386
Dy 1442 TCGCGAGTGTCTCAACCTCAGATCAGCGGAGTACGTCATAGAGCTGACTACTTC 1501
Qy 386 ----- 386
Dy 1502 CCGTCCATCGGCACACGCGCGTATACCCCTCTTACCGCCCAACCCAGACTATAG 1561
Qy 387 -----GlnGlnSerGlnHileSer 392
Dy 1562 ACTGTATGAAGACGCTCAACAGGGGCTGTGATGATGCGAAGCGGAGCACTCGCTG 1621
Qy 393 ThSerleuThrPheGly----- 398
Dy 1622 CGAAACCTCAACATCGGAGCTGTACGCTACAGACAGCCCGGCGCTGTCTACAGC 1681
Qy 399 -----GlnGlyAlaGlnSerProGlyGlnleGlnSerCyAsnArgGlyLys 413
Dy 1682 CAGCCGAGATCCGCGAGACGACAGCTCCCTCGCAGCGGCGGACACTGCCCTTC 1741
Qy 414 GlnProLysValSerAlaGlnGlnProGlySerHileProSerProAlaProArgArgSer 433
Dy 1742 AGCTGAGCTACAGCTTCCACAGCGCGTCTCCCTTACCTTACCTTCCGAGCGGCGCC 1801
Qy 434 ProAlaGlnLysGlnAlaAPRAlaAlaSerAlaProThrglnGlnGlnGln 453
Dy 1802 CTGTGGGCGGCGGTACAGCTGCGGAGCTGACCAATGCGCAGCTGACGCGGAGACTAC 1861
Qy 454 ValValLysAPR-----ArgThrglnGlnSerLysProGlnProGlnProGlnPro 471
Dy 1862 CCGTCTCCCAACATCAGCGGAGCGAGTGAACGCGCACCCCACTTACCCGCCCCC 1921
Qy 472 GlySerleuThrglySerProHileSer---GlnleuSerValAsnSerGlnGly 490
Dy 1922 AGCGCCGCAACAGACGAGCTGTCTCCGCACTTTACATCAGACAGAGCAAC-- 1978
Qy 491 ValAlaProAlaAsnValThrleuSerProAsnleuSerProAPRThrylleGlnAlaSer 510
Dy 1979 -----CCGACCTCATCAGCGCGGCGGTGACACACTCGGTGCAAAAGTTCCAGAGAGAC 2032
Qy 511 ProLeu---lleserProleuLeuAsnAPRAlaCyAsnProArgThr 525
Dy 2033 AGCTTCCCGTGGCGCACTGCTGACAGAGTCAAGAGCCCTTACCC 2080

```

RESULT 3

US-10-505-928-228
 ; Sequence 228, Application US/10505928
 ; Publication No. US20060088532A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ludwig Institute for Cancer Research et al.
 ; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
 ; FILE REFERENCE: 28967/39178
 ; CURRENT APPLICATION NUMBER: US/10/505, 928
 ; CURRENT FILING DATE: 2004-08-27
 ; PRIOR APPLICATION NUMBER: US 60/363, 019
 ; PRIOR FILING DATE: 2002-03-07
 ; NUMBER OF SEQ ID NOS: 866
 ; SOFTWARE: PatentIn 3.2
 ; SEQ ID NO 228
 ; LENGTH: 6469
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-10-505-928-228

Alignment Scores:

Pred. No.:	0.00421	Length:	6469
Score:	176.50	Matches:	160
Percent Similarity:	36.9%	Conservative:	102
Best Local Similarity:	22.5%	Mismatches:	283
Query Match:	3.2%	Indels:	168
DB:	6	Gaps:	32

US-09-555-342b-2 (1-1045) x US-10-505-928-228 (1-6469)

```

Qy 311 GlnHisAlaPhePheArgLeuPheGluGluPro-----LysProLysPro-LysPr 328
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 3436 GAATCGGGTGTGATTTCTTGACCTTGAAGAGCGGGGTAGACCAATGCCAGCATATA 3495

Qy 328 ovalleuPheSerArgLysSerPheArgPheSerGlyArgThrGlnLysGlnValle 348
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 3496 TGTGCTGCTCGAGCGCGCGCT-----AAGCAGGCGCT 3528

Qy 348 uAPRYValLysGluGlyHisLysValGlnPheGluArgLysHisSerLys11 368
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 3529 TGACTGATCCAGAAACAGGT-----GAATTT-----TACCTTCAAC 3567

Qy 368 eHisSerLeaSerLeuAlaSerGlnProThrGlnLeuAlaSerGlnValleuGlu1 388
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 3568 ACATACCTCCACGAGAGACACAGAGAGACTCAGGA-ACGTCTAAAGAAATATGCGG 3626

Qy 388 nSerGlnGlnSerTherSerLeuThrPheGlyGluGly-----AlaGluSerProGly1 406
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 3627 AATTCAAGGTCCTGCCAGCAACAAAGAGAGGAGGAACTTGTATTCACTGCGCG 3686

Qy 406 yGlnSerCysArgArgLysGlyLysGluProLysValSerAlaGly-----GluProGlyse 424
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 3687 ATAGCTTTGT-----GAAAGAGCCCATTCATTCGCCAGCAGATAGAAATGGGTGA 3740

Qy 424 rHisProSerProLarProArgArg--SerProAlaGlyAlaSerGlnAlaaspGlyAla 443
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 3741 CCACGGTGCAAGCACTACAGAGATTTCTCCGAGAGATGGAAAGTACCAATACATCAC 3800

Qy 444 AlaserLarProThrGlnGluGluGluValValLysAspArgThrGlnGlnSerLys 463
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 3801 TGGAGAAAGCCCTTAGAGTCAACAGAGATTAAT-AAAGAC-----3841

Qy 464 ProGlnProProGlnProSerThrGlySerLeuThrGlySerProHisLeuSerGluLeu 483
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 3842 -----CTGGAACCTG 3850

Qy 484 SerValaSerGlnGlyValAlaProAlaAsnValThrLeuSerProAsnLeuSer 503
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 3851 GAT-----ATTATCCAGCAGAC-----3868

Qy 504 ProAspThrLysGlnAlaSerProLeuLieserProLeuLeuAsnAspGlnAlaCysPro 523
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 3869 -----CTTTCGATCGGAGGTCAAG 3889
  
```

```

Qy 524 ArgThrAspAspGluAspGlu---GlyArgArgLysArgPheProThrAspLysAlaTyr 542
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 3890 CTCGGGAGCCCAACACAGATCATGAAGAGACCGGAAGTCACCCCGGAAGAAAGA 3949

Qy 543 PheLeuAlaLysGluValSerThrThrGluArgThr-TyrlleuLysAspLeu---GluVal 561
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 3950 TTTATTATGCTGATCACTCCAGACAGAGAGGCTTATGTATGAGATTTGCTGAGTGC 4009

Qy 562 IleThrSerThrPheGlnSerThrValSerLysGluAspAlaMetProGluAlaLeu--- 580
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 4010 TTAGAGACCTACCTGGGAAATGACAGGTGTGTGAGAGAGATCCCGCTGGGATCTTC 4069

Qy 581 -----LysSerLeuLlePheProAsnPheGluProLysHisLysPheHisThrAsn--- 597
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 4070 AATTAAGAGATATCATCTTTTGCCAAATCCAAAGAGATCTACGATTTCCATTAACAATC 4129

Qy 598 PheLeuLysGluLleGluGlnArgLeuAlaLeuTrpGlu-----GlyArgSerAsnAla 615
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 4130 TTCTCAAGAGCTGAGAGAAATGACGAACTGCTGAGAGATGCGACACTCTTTGTT 4189

Qy 616 GlnIleArgAspTrpGlnArgLleGlyAspValMetLeuLysAsnLleGlnGlyMetLys 635
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 4190 ACCTGGGACAGCAAAATTTCAAGATGATGTCACCTACTGTAAACAAAGCTGATCCAC 4249

Qy 636 HisLeu-----AlaAlaHisLeuTrpLysHisSerGlnAlaLeuGlnAlaLeu 651
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 4250 CAGCTTATCTGAGAGCATGGGCGCACCTTTTGATGATGATACACAGCGGCAGTGTCTG 4309

Qy 652 GluAsnGlyLleLysSerSerArgLysGluGluAsnPheCysArgAspPheGluLeuGln 671
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 4310 GCCAATCTCATCTCTTCC-----4327

Qy 672 LysValCysTrpLeuProLeuAsnThrPheLeuLeuArgProLeuHisArgLeuMetHis 691
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 4328 -----TACCTATTAAAGCTGTCCAAAGAGATCACCAAA 4360

Qy 692 TyrlleGlnValleuGluArgLeuCysLysHisProProSerHisAlaAspPheArg 711
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 4361 TATCAATGCTCTGTAAGAACTTTAACTGCTGTGAAGAGAGGAGGAGCTCAAG 4420

Qy 712 AspCysArgAlaAlaLeuAlaGluLleThrGluMetValAlaGln---LeuHisGlyThr 730
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 4421 GATGGCTGAGGAGTGTGCTCAGTGTCCCAAAAGAGCCCAATGATGCTGATGTCAGC 4480

Qy 731 MetLleLysMetGlnAsnPheGlnLysLeuHisGlnLeuLysAspLeuLleGlyLle 750
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 4481 ATG-----CTGAAAGGTTCCAGAGAACTGAGATGCGAGGGAGTTGATTTCCAG 4534

Qy 751 AspAsnLeuValVal---ProGlyArgGluPheLleArgGluGlySerLeuSerLysLeu 769
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 4535 GATGCTTTCAAGTGTGGAGCCGGAAGTGTGATCCGGAAAGGG-----4579

Qy 770 SerGlyLysGlyLeuGlnGlnArgMetPhePheLeuPheAsnAspValleuLeuTyThr 789
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 4580 -----CGGAGCGGCACTTGTCTTTGAGATCTCTTGTGATTTTAAAG 4624

Qy 790 SerArgGlyLeuThrAlaSerAsnGlnPheLysValHisGlyGlnLeuProLeuTyGly 809
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 4625 AAGAGATCAAGATCTTCAAGACACAGAAATGATTATCAAGAAACAAAGCTTACGACC 4684

Qy 810 MetThrLleGluGluSerGluAspGluTrpGlyValProHisCysLeuThrLeuArgGly 829
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 4685 TCAGAGCTGGGTGTACCGAGCACGCGAGGAGATCCCTGCAAAATTCGCTGTGTCT 4744

Qy 830 GlnArg-----GlnSerLleLeuAlaAlaAspSerArgSerGluMet 844
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 4745 GGGCGACACCCCATCTCAAGCATTAACAGAGCTGAAAGCTTCAACATTAAGAACCAAG 4804

Qy 845 GluLysTrpValGluAspLleGlnMetAlaIleAsp-----LeuAlaGlyLysSer 861
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 4805 CAGAGTGATTCAGAAACATTGAGAAAGTGATTTCAAGAAAGATCTTCACTGAAAGGA 4864
  
```

```

Qy      862  SerSerProAlaPProGluPheLeuAlaSerSerProPro-----AspAenLysSer  878
Db      4865  GCTTTAAAGAGACCCTTCAGCTCCCCCAAAACACAGACCCMAACAGAGAAACAATGTGAAG  4922
Qy      879  ProAerGluAlaIthrAlaIaIaerPngInuSerGlu-----AspAerLeu  893
Db      4935  AGGAGTGGAGTGGAGAGATTTCACACAGCCAGGGGAGATGGAGACGCCAACACACACCATC  4988
Qy      894  Ser---AlaSerArgThrSerLeuGluArgGlnAlaProHlaArgGlyAsnThrMetVal  912
Db      4985  TCCATTGGCTTCAGAGCCCTCCAGAACACAGTGCACAGTGCACAGAGATGGCAACTTGT  5044
Qy      913  HisValCysThrPheIaIaArgAsnThrSerValSerMetValAerPheSerIleAlaValGlu  932
Db      5045  ---CCTCGGTGGCACCT--GGAGACCTGAGATCCTTCTCCATTACGTTAGCG-----  5094
Qy      933  AenGlnLeuSerGlyAsnLeuLeuArgLysPheLysAsnSerAenGlyTTPrGlnLysLeu  952
Db      5095  ---GCCATCTCTGGGACTTGTCTCCAGACCTCACCGGGTTAGCGCGTGGC-----  5139
Qy      953  TrpValValPheThrAsnPheCysLeuPhePheTyrLysSerHisGlnAsp---AsnHis  971
Db      5140  ---ACCGTTGGAGCC-----TCCACACAGACATCCCATAC  5172
Qy      972  ProLeuAlaSerLeuProLeuLeu  979
Db      5173  CCA-----CCTCTCTTG  5184

RESULT 4
US-11-217-529-2747
; Sequence 2747, Application US/11217529
; Publication No. US20060099612A1
; GENERAL INFORMATION:
; APPLICANT: SUNTORY LIMITED
; APPLICANT: NAKMO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: S-38-285
; CURRENT APPLICATION NUMBER: US/11/217,529
; CURRENT FILING DATE: 2005-09-02
; PRIOR APPLICATION NUMBER: US 10/932,182
; PRIOR FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2747
; LENGTH: 2559
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-11-217-529-2747

Alignment Scores:
Pred. NO.:          0.00215          Length:          2559
Score:              173.00           Matches:          114
Percent Similarity: 38.04           Conservative:     83
Best Local Similarity: 22.04         Mismatches:      187
Query Match:        3.2%             Indels:          137
DB:                  7               Gaps:           23

US-09-555-342B-2 (1-1045) x US-11-217-529-2747 (1-2559)
Qy      482  GluLeuSerValAenSerGlnGlyLysValAlaProAlaAenValThrLeuSerProAs  501
Db      681  GACCTCTCAATTGGTCAAAAGTGTGAGGTAGTGAACCGCTAGTAATTCAGTCCAA  740
Qy      501  nLeuSerProAerThrLysGlnAlaSerProLeuLysSerProLeuAenAenArgGlnAl  521
Db      741  TATCTTCCCTCCAGAGCAAAACACAGCAACTTG-----  777
Qy      521  aCyEProArgrThrAerAerGluAerPngInuGlyArgArgLysArgPheProThrAspLys  540

```

Db	778	-----ACTGCAAAAATCAACAGCATCTTCAACAGCAGATCTCTCAAAAAACA	827
Qy	541	---AATyrPhe---IleAlaValSerThrThrglnArgThrTyrLeuLysAs	558
Db	828	TATTAGATGATCTTAAGATATTATCAAAAGAAATTCGTGCCAACAAAGAAAGATACGTTCAATGA	887
Qy	558	PLEuGlnValIleThrSerTrpPheGlnSerThrValSerLysGluAspAlaMetProG1	578
Db	888	TTTAAAGATTTTGGACAAGATATAGCAACAACGCTATTAGACAGTGCACCTGATCAGCTTCGA	947
Qy	578	uAlaLeuLysSerLeuIlePheProAspPheGluProLeuHisLysValPheHisThrAspH	598
Db	948	AGAGCTT---TATATGTTGTTCCCACTTGAGCGACCGCATGATATTTCAAGAAAGATT	1004
Qy	598	eLeuLysGluIleGlnGlnArgLeuAlaLeuTrpGlnGlyArgSerAspAlaGlnIleAr	618
Db	1005	TTTACTATACCTCCGGG---ATCAATGCTCTTAGTGA	103
Qy	618	g---AspTyrGlnArgIleGlyAspValMetLeuLysAsnIleGlnGlyMetLysIle	637
Db	1038	ACCTTCTAAGCAGAGAACCGGAGCTCTTTCATGCAT---	1074
Qy	637	uAlaIleHisLeuTrpLys-----HisSerGluAlaLeuG1	649
Db	1075	---TCCAAACATTTTTCAGGCTTTATAGCCTTGTCATTCGTCAAAATGCGGCATATGA	1133
Qy	649	uAlaLeuGlnAsnGlyIleLysSerSerArgArgLeuGlnLysAsnPheCyAspArgPheG1	669
Db	1134	GTTTCTTACCTCCACTTA-----AACAAATGACATCTGATGA	1177
Qy	669	uLeuGlnLysValCySer-----LeuProLeuAsnThrPheLeuLeuArgProLe	686
Db	1173	ATCGCAGAAGTTCAATATCAACAATAACTAGAAATTCGACAGTCTTCTCTATAGACCGGT	1234
Qy	686	uHisArgLeuMetHisTyrLysGlnValLeuGlnLysGluArgLysValHisLeuProPse	706
Db	1233	GCMAAGGCTTCGACAGTATCTCTATATGATTTAAAGAAATTTGGTGCTTACTAGATGA	1295
Qy	706	rHisAlaAspPheArgAspCyAspAlaAlaLeu-----	717
Db	1293	CAAC---AATACAGAGAACTGGAAGCTGCTTGATATCTTCCAAAGATATGCCAAGAG	1344
Qy	718	-----AlaGlnIleThrGlnMetValAlaGlnLeuHisGlyTh	730
Db	1350	TATCAAGAGAACCAAGAGAAAGACAGAAATCACCAAGTGGTAAAAAACCTTATAGTGATG	1405
Qy	730	rMetIleLysMetGluAsnPheGln---LysLeuHisGluLeu-----Lys	744
Db	1410	AGTGCTAATTTGGAAGGCTATAGCATCTTAAGATTTGGTGAATTTATGTACTTCGATGA	1465
Qy	744	sLysAspLeuIleGlyIleAspAsnLeuValProGlyArgGluPheIleArgLeuG1	764
Db	1470	AGTGTTCAATTTCCACACAAACAATTTCAAGAACCTGAAAAGAGATTC-----	1518
Qy	764	ySerLeuSerLysLeuSerGlyLysGlyLeuGlnGlnArgMetPhePheLeuPheAsnAs	784
Db	1519	-----GAAGTTATCTTTTGAAGAAAATTTATCATATGTGTTTTCAGA	1555
Qy	784	pValLeuLeuTyrThrSerArgGlyLeuThrAlaSerAsnGln---PheLysValHis	802
Db	1560	AGTATAGCTTCAAGAAATTCACATC-ATCACTTAAGTCTCAAGAAAAAATCTTCGCGCTCAG	1618
Qy	802	sGlyGlnLeuProLeuTyrGlyMetThrIleGlnGlnSerGluAsnGluTrpGlyValPr	822
Db	1619	CAT-CAATGTCA-----ACTGGAGGCATGACACACAT	1653
Qy	822	OHisCySerLeuThrLeuArgGlyGlnArgGlnSerIleIleValAlaIleAspArgSe	842
Db	1654	AACAGTCTCAACATTAATTAATCAAGAAGACGACATATATAGT-AGTATATTAATAT	1712
Qy	842	rGluMetGlnLysTrpValGluAspIleGlnMetAlaIleAspLeuAlaGluLysSerSe	862
Db	1713	CCATTTA---TCCTTC	1724

```
QY 862 rSerProAlaProGluPheLeuAlaSerSerProProAspAsnLysSerProAspGluAl 882
DB 1725 TTGCTGACGACGGCGCATTAATAATTCAT-----AT 1757
QY 882 aThrAlaAlaAspGlnGlnSerGluAspAspLeuSerAlaSerArgThrSerLeuGluAr 902
DB 1758 CAACGGCGGTATACATTCATCAACAATTCGTTATCTTCAATTAATATATATAC 1817
QY 902 GGLAlaProHis-----ArgGlyAsnThrMetVal-----HisValCy 915
DB 1818 CAATGAACCTTAACCTGATCTGAGAGGTGAATTATGATATGAATCTTAACCAAAATAT 1877
QY 915 sTrpHisArgSerThrSerValSerMetValAspPheSerIleAlaValGluAsnGlnLe 935
DB 1878 ACCCAAAACACACAGATCATTAACATTTACCTGGGATCTCTTAAGAGCA----- 1929
QY 935 uSerGlyAsnLeuLeuArgLysPheLysAsnSer-----AsnGlyTrp 949
DB 1930 ----GGAATTTCTTTTAAAGTTCAAAAATGAAGAAACGAGATTAATTGG 1977

RESULT 5
US-10-488-619-2736
; Sequence 2736, Application US/10488619
; Publication No. US20060099578A1
; GENERAL INFORMATION:
; APPLICANT: Greenlee, Winner and Sullivan, P.C.
; TITLE OF INVENTION: Human Mitochondrial DNA Polymorphisms, Haplogroups, Associations
; FILE REFERENCE: 98-01 WO
; CURRENT APPLICATION NUMBER: US/10/488, 619
; CURRENT FILING DATE: 2004-03-01
; NUMBER OF SEQ ID NOS: 3040
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2736
; LENGTH: 646
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-488-619-2736

Alignment Scores:
Pred. No.: 0.000659 Length: 646
Score: 169.00 Matches: 47
Percent Similarity: 50.9% Conservative: 38
Best Local Similarity: 28.1% Mismatches: 72
Query Match: 3.1% Indels: 10
DB: 6 Gaps: 4

US-09-555-342B-2 (1-1045) x US-10-488-619-2736 (1-646)
QY 660 ArgPheGluAsnPheCyAspAspPheGlu-----LeuGlnLysValCyAsp 675
DB 32 AAGGTTCAGCAAGTACGTACTTCTTGAAGCCCTGCGCGCTGCTGCAAGGATATTGAC 91
QY 676 LeuProLeuAsnThrPheLeuLeuArgProLeuHisArgLeuMetHisArgLysGlnVal 695
DB 92 ATCTCGGTGATGCTTTCTGCTGACCCCGGTGCAAGAAATCTGCAAGTATCCCTCAG 151
QY 696 LeuGluArgLeuCyAspLysHisArgProProSerHisAlaAspPheArgAspArgAla 715
DB 152 CTGGCGGAGCTGCTCAAGTACACACACCCTGACAGGAGCTTTAAGATGTTGAAGCT 211
QY 716 AlaLeuAlaGluIleThrGluMetValAlaGlnLeuHisArgLysThrMetIleLysMetGlu 735
DB 212 GCTTTCACCGCATGAAGAAATGTGCGCCAGCTCATCAACGAAACGAAAGACTTGAA 271
QY 736 AsnPheGlnLysLeuHisArgLysLys-----LysAspLeuIleGlyIleAspAsn 752
DB 272 AACATCGACAAGATGTGCTGACGACAGCTCCATAGAGAGCTGGAGGGG---GAGAC 328
QY 753 LeuValValProGlyArgGluPheHisArgLeuGlySerLeuSerLysLeuSerGlyLys 772
DB 329 CTCTGTGTACGAGACTCGAATCATCTCGGGGAGGCTGAGCCGCTGACACACAGCT 388
```

```
QY 773 GlyLeu-----GlnGlnArgMetPhePheLeuPheAsnAspValLeuLeuTrpThrSer 790
DB 389 CAACCGCAGAGTACGACGAGAAATGTTTCTCTTTCACCCGACGCTCATCTACTGAAA 448
QY 791 ArgGlyLeuThrAlaSerAsnGlnPheLysValHisGlyGlnLeuProLeuTrpGlyMet 810
DB 449 AAGGATCTGCTGCGCGGAGGTGATGTAACAAGGCGCGCTGACATGATGACCTG 508
QY 811 ThrIleGlnLysSerGluAsp 817
DB 509 GAGGTGTGATGATGTGAAGAT 529

RESULT 6
US-11-217-529-77046
; Sequence 77046, Application US/11217529
; Publication No. US20060099612A1
; GENERAL INFORMATION:
; APPLICANT: SUNTORY LIMITED
; APPLICANT: NAKAO, YOSHIMIRO
; APPLICANT: NAKAMURA, NORIHISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: S-38-285
; CURRENT APPLICATION NUMBER: US/11/217, 529
; CURRENT FILING DATE: 2005-09-02
; PRIOR APPLICATION NUMBER: US 10/932,182
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 77046
; LENGTH: 4119
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-11-217-529-77046

Alignment Scores:
Pred. No.: 0.0653 Length: 4119
Score: 155.50 Matches: 190
Percent Similarity: 37.8% Conservative: 140
Best Local Similarity: 21.8% Mismatches: 321
Query Match: 2.8% Indels: 223
DB: 7 Gaps: 49

US-09-555-342B-2 (1-1045) x US-11-217-529-77046 (1-4119)
QY 22 IleSerThrLeuGlnLysArgGlnLysProProThrProSerGlyLysLeuValSer 41
DB 1111 ATCAATTTTATGAACTGGCTTAAGATGATCTTAATAGTGTGATCATGAGAACAA 1170
QY 42 IleLys-----IleGlnMetLeuAsp 48
DB 1171 ATTAACCAATTCAATCCAAAGATCTGCATATTGACTCGATGATAAAACCTACCAAC 1230
QY 49 AspThrGlnLysAlaPheGlnValProGlnArgAlaProGlyLysValLeuLeuAspAla 68
DB 1231 GATACGAATTCGCTTAC-----CCAAACAAAGATATGCCAAGATAGAAAGGAGCT 1284
QY 69 ValCys---AsnHisLeuAsnLeuValGlnGlyAspTrpPheGlyLeuGlnPheProAsp 87
DB 1285 CTATGTCTTAATAATGCTGCTTATTATGAAGCAAAAGATAGTCCCGTAGAGGACCAATA 1344
QY 88 HisLysLysIle---ThrValTrp-----LeuAspLeuLeuLysProIleValLysGln 104
DB 1345 AACGAGATTATTAATACGCTGTGAGAAATCTTAGACTTCAAAACCATATCGAATCT 1404
QY 105 IleArgArgProLys-----HisValValValLysPheValVal 117
DB 1405 ATTAACTATTAAATATGATTAATCTTTGCTGTTTATATCTCATATAGATTCCTTCAAGT 1464
QY 118 LysPhePheProPro---AspHisThrGlnLeuGlnGlnLeuThrArgTrpLeuPhe 136
```


Db 1465 TCAAGATCCCTTTTGATGAACCTTAGAGCCGGAAGAAATNGGATTAATCTTTT 1524
 Qy 137 AAlaLeuGlnValLyvGln-----AapLeuAlGlnGlyArg 148
 Db 1525 CAAAGATCTGTAAATGATTTGGATTCTGTACAACTGATGAAATGGTAAGAGCG 1584
 Qy 149 LeuThr---CybaAnaPThrSerAlaAlaLeuLeuIleSerHisIleValGlnSerGlu 167
 Db 1585 GTCCACAGAAATTGATGATCTTAATGCTTAG-----ATTTCATCTTG-----AATGAA 1632
 Qy 168 IleGlyAspPheAspGlnAlaLeuAspArgGlnHisLeu---AlaLyvAsnLyvGlyTle 186
 Db 1633 AAATTAATTAATTAAGAAATCATGATTAAGATCATCTTAATGCAAAATCATGAGAAAGT 1692
 Qy 187 ProGlnGlnMetAlaLeuGlnAspLyvIleValGlnPheHisIleValGln 206
 Db 1693 GAATCTTTGATATCTCGAATAACCAAGAAATTGAA----- 1728
 Qy 207 ThrProAlaGluSerAspPheGlnLeuLeuGlnIleAlaArgGlnLeuMetGlyTrol 226
 Db 1729 -----AATCGAAATCGAGTTGAAGCAACCAAGAAAGCTAGACCA----- 1773
 Qy 227 IleArgLeuHisProAlaLyvAspArgGlnGlyThrLyvIle-----AsnLeuAla 243
 Db 1774 ATTACCAAGCATCAAGGTTATATGACACGACCTTATGGCAAGTAGCACTTGAGT 1833
 Qy 244 ValAlaAsnThr-----GlyIleLeuValPheGlnGlyPheThr 256
 Db 1834 ATTCCCGCTCGAATAATTAATAATATGATCATGAAATATCATTTTCCAAATTTTGACA 1893
 Qy 257 LyvIleAsnAlaPheAsnThrAlaLyvAlaArgLyvLeuSerPhe---LyvArgLyvArg 275
 Db 1894 AAA-----AAGCAACAGCAGCAGCAAAATTTCTGCCAAACCGTTCAACT 1941
 Qy 276 PheLeuIleLyvLeuArg-----ProAspAlaAsnSer 286
 Db 1942 AGCTACTATAAAAGTAAAGAGTTACGTCCTTATCATCATTTTAATGATGAGCAAAATAC 2001
 Qy 287 AlaTyrGlnMetThrLeuGlnPheLeuMetAlaSerArgAspPheCyvLyvSerPheTyr 306
 Db 2002 GAAATAGTCCCA----- 2016
 Qy 307 LyvIleCyvAlaGlnHisIleAlaPhePheArgLeuPheGlnGlyProLyvProLyvPro 326
 Db 2017 -----AATGAAATCCGAAGATTAATCA 2037
 Qy 327 LyvProValLeuPheSerArgGly---SerSerPheArgPheSerGlyArgThrGlnLyv 345
 Db 2038 AAAGACTCATTAATTTCAAAAGATCTACGTCACATTAATTTAATATCCCTCTATGAAA 2097
 Qy 346 GlnValLeuAspTyrValLyvGlnGlyGlyHisLyvLyvValGlnPheGlnArgLyvHis 365
 Db 2098 AATATTCTATAATAG-----CAAAATGCTCTTAATTCGATACTT 2139
 Qy 366 SerLyvIleHisSerIleArgSerLeuAlaSerGlnProThrGlnLeuAsnSerGlnVal 385
 Db 2140 TCGAGTTGGAATTTTCAAAATAGCTAGAACACAACTTAATTAACAATGCTCCCGTG 2199
 Qy 386 LeuGlnGlnSerGlnGlnSerThrSerLeuThrPheGlnGlnGlyAlaGluSerProGly 405
 Db 2200 TTGTCTTCGGTTCTTCTACCAAACTT-----TTCCTCGACTTCC 2244
 Qy 406 GlyGlnSerCyvArgArgGly-----LyvGlnProLyvVal 417
 Db 2245 TCAAGATGCTTGAATAATGGAATCCAACTAGTACCTGAAGTTGTTAAACTACCTCACTT 2304
 Qy 418 SerAlaGlnGlyProGlySerHisProSerProAlaProArgArg---SerProAlaGly 436
 Db 2305 CCGCGCGCGCGCTCCACCTCCCGCTCTCCTCACTTCACACAGCTCTTTGAGTGAAGCA 2364
 Qy 437 AsnLyvGlnAlaAspGly-----AlaAlaSerAlaProThrGlnGlnGlnGlu 452

Db 2265 GAAGCTAAACCGAGATGGGTGTTCCGTATATGACAGCCTGTCCG---CCACCCCTTCCA 2421
 Qy 453 GlnValValLyvAspArgThrGlnGlnSerLyvProGlnProPheGlnProSerThrGly 472
 Db 2422 GATCTATTAAATTAAGCTAAACCTTGCGCGCTTCACACACACCG----- 2466
 Qy 473 SerLeuThrGlySerProHisIleSerGlnLeuSerValAsnSerGlnGlyValAla 492
 Db 2467 -----CCTTACCCGAGCTTGTTCGATGAACAAAGC----- 2499
 Qy 493 ProAlaAsnValThrLeuSerProAsnLeuSerProAspThrLyvGlnAlaSerProLeu 512
 Db 2500 CCCCTCAATACAGATTAGT-----ACTCTCGAGCTCCACTTTA 2541
 Qy 513 IleSerProLeuLeuAsnAspGlnAlaCyv-----ProArgThrAspArgGlnAsp 529
 Db 2542 CCAATGCTCTCTTCTCTCTCCCTCAGTGTCTATCAATCCAAAGCAACCGAT----- 2595
 Qy 530 GlnGlyArgArgLyvAspPheProThrAspLyvAlaTyrPheIleAlaLyvGlnValSer 549
 Db 2596 -----TTAAACCACTCCACTGAAGAGCA-----TTGAACCAATCCAC 2637
 Qy 550 ThrThrGlnArgThrTyrLeuLyvAspLeuGlnValIleThrSerTyrPheGlnSerThr 569
 Db 2638 TGGGATTAAGGAGAGATATCAAGAC-----ACACTTGG---GAAGACACG 2682
 Qy 570 ValSerLyvGlnAspAlaMetProGlnAlaLeuLyvSerLeuIlePheProAsnPheGln 589
 Db 2683 TTTCACGCCCAAGAAACAAACAAAGAAATTAACAATGATGATATCTCTCAAGTTGAA 2742
 Qy 590 ProLeuHisLyvPheHis-----ThrAsnPheLeu-----LyvGlnIleGlnGlnArg 605
 Db 2743 GATATTTTAAAGATGAAGAAAGTCCGACTGAAGATTGTATAAAGCAAGCAAGTCTTCA 2802
 Qy 606 LeuAlaLeuTyr-----GlnGlyArgSerAsnAlaGlnIle-ArgAspTyrGlnArgG1 623
 Db 2803 ATTGCTTGCTTCAAAACATAGCAACTCTTCCATGACATGAGAAATTTCAATTTCTTA 2862
 Qy 623 eGlyAspValMetLeuLyvAsnIleGlnGlyMetLyvHisIleValAlaHisIleTyrLy 643
 Db 2863 TCAAGAGATTGGCTCGACGATTTGGTATTAATTTA-CACATGTTTCCCAATATATCGA 2921
 Qy 643 HisSerGlnAlaLeuGlnAlaLeu-----GlnAsnGlyTle---LyvSerSerArgArg 660
 Db 2922 TATGGAATTTGTATGAAGATTAATGAACGTGTATGAAGACATGTTCAAAATGTTAACT 2981
 Qy 660 GlnGlnAsnPheCyvArgAsp-----PheGlnLeu-GlnLyvValCyvTyrLeuProL 678
 Db 2982 ACTGAATTTTGTGTAAGAAGAGTTAGTAATTAATACCAAAAGTATGCTTAATAGTA 3041
 Qy 678 eHisThrPheLeuLeuArgProLeuHisArgLeu-----MetHisT 692
 Db 3042 CGAGC-----CATATTCACAGGGTAAGATGTAAAGCAGTAAGTACTT 3086
 Qy 692 YrlvsglnValLeuGlu----- 697
 Db 3087 ACAAGAGCTGACGAATATTTTGAAGTTGTGTATCAATTAAGATTTATTTGGAATGT 3146
 Qy 698 -----ArgLeuCyvLyvHisIleHisProProSerHisIleAspPheArgAspCyvArg 715
 Db 3147 AAGATCAAAAGCTGCTGATTAATTAAGAAGAGTTAGTAATTAATCCAAAGTATGCTTAATAGTA 3205
 Qy 715 AlaAlaLeuAlaGlnIleThrGlnMetValAlaGlnLeuHisGlyThrMetIleLyvMetG 735
 Db 3206 TCAAGTTACAAAATTAATGATGATGCAATTTCAACACTTAAT-----CGTTCAAC 3253
 Qy 735 LyvAsnPheGlnLyvLeuHisGlnLeuLyvAspArgLeuIleGly-----IleAspAsnL 753
 Db 3254 CTAAATTTAAAGTTGATGTTTATTTATTAACAAA---ATAGCAATCATATGAATAAAA 3310
 Qy 753 euValValProGlyArgGlnPheIleArgLeuGlySerLeuSerLyvLeuSer-----G 771
 Db 3311 GAATGTGTAAGGT-----ATCAAAATTGAAGTATGACGTAAACTTGCCCTTGCTCA 3361

Qy 868 PheLeuAlaSerSerProProAspAsnLysSerProAspGluAlaThrAlaAlaAspGln 887
Db 1402 ---CTCAAGACGACATCTTGAGAGAGCGCGCCAAATGAGATGAGAGCGCGTGG 1458
Qy 888 GluSerGluAsp-----AspLeuSerAlaSerArgThrSerLeuGlu 901
Db 1459 CAGATCAAGACACCGGTGACATGATGCTGGAGAGCCAAACAGAGAG 1506
RESULT 8
US-10-505-928-311
Sequence 311, Application US/10505928
Publication No. US20060088532a1
GENERAL INFORMATION:
APPLICANT: Ludwig Institute for Cancer Research et al.
TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
FILE REFERENCE: 28967/39178
CURRENT APPLICATION NUMBER: US/10/505,928
CURRENT FILING DATE: 2004-08-27
PRIOR APPLICATION NUMBER: US 60/363,019
PRIOR FILING DATE: 2002-03-07
NUMBER OF SEQ ID NOS: 866
SOFTWARE: PatentIn 3.2
SEQ ID NO 311
LENGTH: 7465
TYPE: DNA
ORGANISM: Homo sapiens
US-10-505-928-311
Alignment Scores:
Pred. No.: 0.25 Length: 7465
Score: 152.00 Matches: 201
Percent Similarity: 31.08 Conservative: 118
Best Local Similarity: 19.54 Mismatches: 339
Query Match: 2.84 Indels: 373
DB: 6 Gaps: 47
US-09-555-342B-2 (1-1045) x US-10-505-928-311 (1-7465)
Qy 152 AsnAspThrSerAlaAlaLeuLeuLysSerHisLeuAlaGlnSerGluLysLeuAspPhe 171
Db 1444 AATGAGCAGCTGACAGATTTCTTGTGCGGACGCTTCAAGCTGAGACAGAGGAATAT 1703
Qy 172 Asp---GluAlaLeuAspArgGluHisLeuAlaLysAsnLysTrpLysProGlnGlnAsp 190
Db 1704 GACCTGAGAGACATTGACTGCTGCACATC-----GAGTTCACTGACACACGAGAT 1754
Qy 191 AlaLeuGluAspLysLeuAlaGluPheHisLeuHisLeuGlnThrProAlaGlu 210
Db 1755 GCCCTG---GACATGATTGCCAACAAGCCCATGAAACATCATCTCCCTCATGATGAGAG 1811
Qy 211 SerAspPhe-----GlnLeuLeuGluLysLeuAlaArgArgLeuGlnMet 224
Db 1812 AGCAAGTTCCCAAGAGGACACAGACACACATGTTACACAGCTGAATCTCCACAGACAAAG 1871
Qy 225 TyrGluLysLeuGluHisProAlaLysAspArgGluGlyThrLysHisLeuAlaVal 244
Db 1872 CTCAAGCCCAACTACATCCCTCCCAAGAAACAACATGAGACCGAGTTGGCATC----- 1935
Qy 245 AlaAsn-----ThrGlyLysLeuValPhe-----GlnGlyPheThrLysHis 259
Db 1926 ---AACCATTTTGCAGGCACTGCTCACTATGAGACCCAAAGCTTCTCGAAGAAAGACGA 1982
Qy 260 ---AlaPheAsnTrpAlaLysValArgLysLeuSerPheLysArgLysPheLeu--- 277
Db 1983 GACACCCCTGATGGGACATATCCAGCTGGTCCACTCTCCACAGAAACAAGTTATCAAG 2042
Qy 278 -----CAGATCTTCCAGGCCGATGTCGCATGGGCGCCGACAGACCAAGAGCGTCCCACTT 2102
Db 2043 CAGATCTTCCAGGCCGATGTCGCATGGGCGCCGACAGACCAAGAGCGTCCCACTT 2102
Qy 285 AsnSerAlaTrpGlnAspThrLeuGluPheLeuMetAlaSerArgAspPheCysLeuSer 304
Db 285 AsnSerAlaTrpGlnAspThrLeuGluPheLeuMetAlaSerArgAspPheCysLeuSer 304

Db 2103 AGCAGCCAGTTCAAGCGGTCACTGAGCTGCTGATGCGCACGCTGGTGCCTGCAGCCC 2162
Qy 305 PheTrpLysLeuCysValGluHisHisAlaPhePheArgLeuPheGluGluProLysPro 324
Db 2163 TTCTTTGTGCATCATCAACAGCCCAATGATGTC----- 2195
Qy 325 LysProLysProValLeuPheSerArgGly-----SerSerPheArgPheSerGly 341
Db 2196 ---AAGAGCCCATGCTGTTTGCACCGGACCTTGCTGCGGCCAGCTGCGGTATCTCAGA 2252
Qy 342 ArgThrGlnLys-----GlnValLeuAsp 349
Db 2253 ATGATGAGAGACCATCCGAATCCGCGAGCTGCTAACCACCTCCGTAACGCTTGAGAG 2312
Qy 350 TyrVal-----LysGluGly--- 354
Db 2313 TTTGTGAGCGGTACCGTGTGCTGTCGACAGTGTGAAGCCGCGCTTCAAGACGGCGAC 2372
Qy 354 ----- 354
Db 2273 CTCGCGGAGACTTCCAGCGCATGCTGAGGCTGCTGGGACCCACAGATGACTGGCAG 2432
Qy 355 ---GlyHisLysValGlnPheGluArgLysHisSerLysHisSerLeuArg-Se 373
Db 2433 ATAGCAAAACCAAGATCTTTCTGAGAGGACACATGACATGCTGCGAAGTGAGCGG 2492
Qy 373 rLeuAlaSerGlnProThrGluLeuAsnSerGluValLeuGlnGlnSerGlnSerThr 393
Db 2493 GACAAAGCATACACCGACAGATCATCTCTTCAGAAAGTCATCCGGGAGTTCAAAAGC 2552
Qy 393 rSerLeuThrPhe-----GlyGluGlyAlaGlu 402
Db 2553 AGCTTAACCTTTCTGAAAGCTGAGAGACGCTGCCACATGATCCAGAGGCACTGGCGGGT 2612
Qy 402 userProGlyGlyGlnSer-----CysArg----- 410
Db 2613 CACAACTGTAGGAAGAACTACGCGGTGATGCGTGGGCTTCTGCGGCTGCAAGCCCTG 2672
Qy 411 -----ArgGlyLys-GluProLysValSerAlaGlyGluProGlySerHis-----P 426
Db 2673 CACCGCTCCCGAGCTGCACACAGATACCGCTGCGCGCGACCATCATCAGATTC 2732
Qy 426 roserProAlaProArgArgSer----- 433
Db 2733 CAGCGCCGCTGCGCGCTATCTGTGCGCAAGGCTTCCGCCACCGCTTGGGCTGTG 2792
Qy 433 ----- 433
Db 2793 CTCAACGTCAGGCTTAAGCCCGGCGATGTCGCCGACGCTGCACCAACGCTCAGG 2852
Qy 434 -----ProAlaGlyAsnLysGlnAlaAspGlyAla-----AlaS 445
Db 2853 GCTGATATCTGTGCGCTTCGAGCTGAGAAATATGCGGCTGCGGAGAGAAAGAACTT 2912
Qy 445 exrAlaProThrGluGluGlnGluValValLysAspArgThrGlnGlnSerLysProG 465
Db 2913 CGAAGAGATGACCGCCAAAGAGCCCAAGAGAGGCGCCAGCCAGCATCAGAGAGCC 2972
Qy 465 InProProGlnProSerThrGlySerLeuThrGlySerProHisLeuSerGluLeuSerV 485
Db 2973 CTGCGCCAGCGCTGCTGAGAGACGCTGAGCGGAGCT----- 3010
Qy 485 alaAsnSerGlnGlyLys-----ValAlaProAlaAsnValThrLeuSerP 500
Db 3011 --GAAGAGAGAGAGGCGCGCTGCGGAGAGAGAGAGCTCTCGAGCAGATGAGAAAGGCC 3068
Qy 500 roAsnLeuSerProAspThrLysGlnAla-SerProLeuLeuSer----- 514
Db 3069 CG-----CATGAGCCTGTCAATCACTCAACATGTGTGACCAAGATGTTGGCTTC 3119
Qy 515 -----ProLeuLeuAsnAspGlnAlaCysProArgThrAspAsp 527
Db 3120 CTGGAGACTTCAGGTGCTGTCGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3179

Dh 823 ACAACAGATACAAAGTATGTAGTGAATTCGTGATTAAGAAAGTTGAAGGTAAGAAATCCCTTCT 882
Qy 392 SerTherSerLeuThr-----Phe 397
Dh 883 ACCGACATCATCGAGCTTCAACCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 942
Qy 398 G1AG1uG1uA1aG1uSerProG1uG1uInsEr---CyArGArG1uG1uG1uProLys 416
Dh 943 GGTAAAGTTATATCAATCAACGATTAATGAATAAACAAGAAACAAATGCTGCTGCTGCT 1002
Qy 417 ValSer1aG1uG1u-----ProG1uSer1aProSerPro1aProArg 431
Dh 1003 GCTGCTCTGGAGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1062
Qy 432 ArgSerPro1aG1u---Aen1uG1uA1aSerG1uA1aSer1aProThG1uG1u 450
Dh 1063 CCTCCCTCTAATGATTAATTTGGCCAAATATATGAGCA----- 1101
Qy 451 G1uG1uG1uVal1aLysArGArgThG1uG1uInsEr1uSerProG1uProProG1uProSer 470
Dh 1102 -----ACTCCTCCACCAACCTTTACCTTTT 1128
Qy 471 ThrG1uSerLeuThrG1uSerProH1aLeuSerG1uLeuSerValaSerG1uG1u 490
Dh 1129 GTATTAATCTTCAACTGGTGA-----GGTGA 1155
Qy 491 Val1aPro1aAenVal1aThrLeuSerProAenLeuSerProArG1uG1uA1aSer 510
Dh 1156 ATTCACCACTGCT-----CACCCGATGATGCGACGCTTCCAAATGAAATCA 1200
Qy 511 ProLeu1eSerProLeuLeuAenAer-----G1uA1aCyArProArg--- 524
Dh 1201 GCCCTCACATCTCCGTGCTTCTTCAGTCCACCATCGTATTGTAATAATATCCAGTCCC 1260
Qy 525 -----ThArPArPArg1u----- 528
Dh 1261 CACAAAAATTAAAGCAACGTCATTTGGGAAAAAATAGACTCTAGTAATTCATTTGG 1320
Qy 529 AArG1uG1uArGArg1uArGArgPheProThArP----- 539
Dh 1321 GGAATCAGATGAAGCTGAGAAATTTGCGAGATTTGTAATGAAAAAGGTCCTTGCAT 1380
Qy 540 ---LysA1a1uYrPhe1a1aLysG1uValaSerThrThG1uArG1uThr1uYrLeuLysArP 558
Dh 1381 TTAAAGAAAGCCTTGTGCTGCTAGGAAATCAATCTTTAGCCACTAAAGCAAGAAAGAT 1440
Qy 559 LeuG1uVal1a1eThr-----SerTrp 565
Dh 1441 TTACAGAAAGTAACTTTCTTATCCAGAGACATTTCTCAACGTTTGGATCACTTGAT 1500
Qy 566 PheG1uSerThrValaSer1uG1uArP1aMerProG1uA1aLeuLys----- 581
Dh 1501 ATGTACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTT 1560
Qy 582 -----SerLeu1e---PheProAenPheG1uProLeuH1aLysPheH1s 595
Dh 1561 TTTTTCAAACACCTAGTGTGTGAATTTTATCAAAACCTGAGATCAACAGATATCT 1620
Qy 596 ThrAenPheLeuLysG1u1eG1uG1uArG1uA1aLeuThrG1uG1uArG1uArG1uSerAenA1a 615
Dh 1621 GTCAATTTAGCTAAACATAGCCCAATTTCTACCCATTTGGAAAGGCTTAAGGATCTT 1680
Qy 616 G1u-----1LeArGArP1uG1uA1aG1uG1uG1uG1uG1uG1uG1uG1uG1uG1uG1uG1u 627
Dh 1681 GAAGATGCAAAACCTCCAGAAAAAGATCCGATGATTCACAAAGGCGCATCAATTTAC 1740
Qy 628 LeuLys-----Aen1eG1uG1u----- 633
Dh 1741 CTGCAATTTAATGTTATCTGAAATCTGATCGGGGCTTCTGTAAGAGCTTTAAGGTA 1800
Qy 634 -----MetLysH1aLeuA1aA1aH1aLysLeuThrLysH1aSerG1u 646
Dh 1801 GTTACCTCATATGAAGAAAGATATATAGACTGCTGGGGAAGTTGAGAAAGTTGAGAGA 1860

Qy 647 A1aLeuG1uA1aLeuG1u-----AenG1uLysSerSerArGArg 660
Dh 1861 GCTGTGCTGCGCTTCAAGAGTCAAGATTAATTAACGTAACGTTTCAACGTTATTTAGT 1920
Qy 661 LeuG1uAenPheCyArGArPheG1uG1u-----LysValaCyArTrLeu 676
Dh 1921 GTGCGTAATCTTTAAGACATACCTCAAGCAAGCAAGGCTTAATCTTGACATTTA 1980
Qy 677 ProLeuAenThrPheLeuLysArGProLeuH1aArgLeuMetH1sTrLysG1uA1aLeu 696
Dh 1981 CAAGGCTGACATTT---ATCAAGATATACCAATATGATGATCTTTAAATTAATGATGC 2037
Qy 697 G1uArGLeuCyArLysH1sH1sProProSerH1aA1aArPhe----- 710
Dh 2038 GAAAAATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2097
Qy 711 -----ArGArPArG1aA1aLeu 717
Dh 2098 GTACTAGATGCTGCAAGCTCTCATTAAGAACATTAATTAATGACTGTAAGATTTAGT 2157
Qy 718 A1aG1u1eThrG1uMetVala1aG1uLeuH1sG1uThrMet1eLysMetG1uAenPhe 737
Dh 2158 CAATCAATTTGTAATGATGAGCGCTCAGTTGA-----ATGGTAATTTAAGTACTCT 2211
Qy 738 G1uLysLeuH1sG1uLeuLysArPLeu1eG1uLys----- 750
Dh 2212 TCGAAGTTCAATTCATTCATTAAGTAAAGTTCTGATTAATAAATTCCTGCTGATCAACAGAGCA 2271
Qy 751 -----AsPArLeuVala1aProG1uArG1uPhe1eArG 762
Dh 2272 AGAAGAAAGGAGATTTATTAAGAAAGAAAGTTAAGTTAATCATATGAGTTTGAAGT 2331
Qy 763 Leu---G1ySerLeuSerLysLeuSerG1uLysG1uLeuG1uA1aG1uMetPhePhe--- 780
Dh 2332 TTATGCAAAACATATGCTGAGATTTCTGATTAAGTTCCAAAGATTTCAATTTTAAAG 2391
Qy 781 -----LeuPheAenArPValaLeuLeuThrThSerArG1uLeuThrA1aSer 796
Dh 2392 AAATTCGACATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2451
Qy 797 AenG1uPheLysValaH1sG1uG1uLeuProLeuThrG1uMetThr1eG1uG1uSerG1u 816
Dh 2452 GAAGAAAGAAAGCTTTTGAAGAG-----CATAGAAATTCATGAGAAAGCAACAA 2502
Qy 817 AArG1uThrG1uValaProH1sCyArLeuThrLeuArg1uG1uArG1uInsEr1e1eVala 836
Dh 2503 AAAAG-----GCAACAAGAAAGAAAGCAACATGAGAGTCT 2538
Qy 837 A1aA1aSerSerArGArgSerG1uMetG1uLysTrpValaG1uArP1eG1uMetA1a1eArP 856
Dh 2539 AGAGCTTACCCGAAAGAAATGAGAAAGAAAGAAAGAAAGCAAGCTGCTTTATGAT 2598
Qy 857 ---LeuA1aG1u-----LysSerSerSerPro1aProG1uPheLeuA1aSerSer--- 872
Dh 2599 AAATTAATTAACAGCTGAAAGAAATGACAGGCCCTTAAGATATCTTGCTGCTGAGA 2658
Qy 872 ----- 872
Dh 2659 AAAAGGCTGCTGTTAGAAAGAAATATCTTTGCAAAAGATTAATTCGAGCAAGTACTC 2718
Qy 873 -----ProProAenLysArPArP 880
Dh 2719 AACGACTGGAATTAAGAAAGTGAATCTATTTCTACTCTCCGAGGCTCAACCCCTGAGC 2778
Qy 881 G1uA1aThrA1aA1aArG1uG1uSerG1uArPArPLeuSerA1a-----SerArg 897
Dh 2779 ACGGATGCAAGATTCATGCTGAATACCTAAGCGGCTTGCCAGAGAGCTTTATGAGC 2838
Qy 898 ThrSerLeuG1uArG1uA1aProH1sArG1uA1aThrMet 911
Dh 2839 ACTTCAGAGATTTTACCATCAACATCAAAAGCATTCGATTTG 2880


```

Qy 459 ThrGlnGlnSerLysProGlnProGlnProSerThrGlySerLeuThrGlySerPro 478
Db 3871 CTTCCACCTGCGAGGCTCTGCGAGACAGACCCAGCT-----TCTCCA 3912
Qy 479 -HisLeuSerGluLeuSerValIAsnSerGlnGlyValAlaProAlaAsnValThrLe 498
Db 3913 GGGCCTCAGTAAATACATGACAGTACAGCCCTCTGGGACCCACCTGCCGGGGACCGT 3972
Qy 498 uSerProAsn-----LeuSerProAsp-----Th 506
Db 3973 CTTCCACAGAGGGAGCTTTCTCAGACCTTCATCTGCTCAAGCCCCGACAGGATGGAC 4032
Qy 506 rLyGlnAlaSerProLeuIleSerProLeuLeuAsnArgAlaCysPro----- 523
Db 4033 CAAGCCATGAAAGCCAGAGGAGCCCACTTCTTGTAATATATACAGACCCCTTTGGAA 4092
Qy 524 -----ArgThrAspAspG1 528
Db 4093 CTTCTTCACACCCCTAAACTGTTGATATACACTGGCTTTCTAGAGCTCAGACCTTT 4152
Qy 528 uArgGluGlyArgArg-----LysArgPheProThrAspLysAlaTyrPheIleAlaLys 546
Db 4153 GGACTCAGAGAACCAAGAGCCCCCAAAAGATCCCTAGACCCA-----GTGTGACG 4203
Qy 546 eGluValSerThrThrGluArgThrTyrLeuLysAspLeuGluValIleThrSerTrpPh 566
Db 4204 TGGGCTTCAGCTGCCAAGGAGGAGAGCTGCCACCAATATGACAAACCTTAGGTCA 4263
Qy 566 eGlnSerThrLysGluAspAlaMetProGluAlaLeuLysSerLeuIlePhePr 586
Db 4264 CTGGGTGACATCATGACGACCAAGAA-----AAGAGACAGCAGCTGCA 4305
Qy 586 AsnPheGluProLeuLysLysPheHisThrAsnPheLeuLysGluIleGluGlnArgLe 606
Db 4306 GGTGTTTGAACCCACTGGCCAAA-----ACATGACT 4335
Qy 606 uAlaLeuTrpGluGlyArgSerAsnAlaGlnIleArgAspTyrGlnArgIleGluAspVa 626
Db 4336 GAGCAGCTTGAAGGCTGCAGTCTTA-----GAATG 4368
Qy 626 IMetLeuLysAsnIleGlnGlyMetLysIleuAlaAlaHisLeuTrpLysHisSerG1 646
Db 4369 CATACCTTCCTCCCTAGACATCCCTCCACAGAAAGACATGATTTAAAGGACACT 4428
Qy 646 uAlaLeuGluAlaLeuGluAsnGlyTLeuSerSerArgArgLeuGluAsnPheCysArg 666
Db 4429 GGCC-----AAAACGTTTGTCATCTGTCACTCTCGT-- 4462
Qy 666 AspPheGluLeuGlnLysValCysTyrLeuProLeuAsnThrPheLeuLeuArgProLe 686
Db 4463 -----AGTTTACAAATGCTGT----- 4480
Qy 686 uHisArgLeuMetHisTyrLysGlnValLeuGluArgLeuCysLysHisAspProSe 706
Db 4481 -----CTTTATTCATTAAGATGTTACTCAG-----CCACCAA 4515
Qy 706 rHisAlaAspPheArgAspCys 713
Db 4516 ATATATTTCACTCAAGGCTGT 4537

```

```

; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 189
; LENGTH: 7268
; TYPE: DNA
; ORGANISM: Homo sapiens - Base pairs nucleotide sequence
US-10-559-415-189

Alignment Scores:
Pred. No.: 1.63 Length: 7268
Score: 140.00 Matches: 114
Percent Similarity: 35.1% Conservative: 79
Best Local Similarity: 20.7% Mismatches: 202
Query Match: 2.6% Indels: 155
DB: 6 Gaps: 21

US-09-555-342B-2 (1-1045) x US-10-559-415-189 (1-7268)
Qy 405 GlyGlyGlnSerCysArgArgGlyLysGluProLysValSerAlaGlyGluProGlySer 424
Db 3 GGGCGGAGG-----CGCAGAGGGAGAGCTGCTGTAAGCTGCTTGGCCAGAGAGCG 56
Qy 425 HisProSerProAlaProArgArgSerProAlaGlnLysLysGlnAlaAspGlyAlaAla 444
Db 57 GGAACCACTTCCA-----GAACTCGGGGGCGGCTGGAGAAAGACGAGCCCTCA 104
Qy 445 SerAlaProThrGluGluGluGluValValLysAspArgThrGlnGlnSerLysPhe 464
Db 105 CTTCTGCCCATCTCTTACTGAAAGCGGTGCT----- 138
Qy 464 oGlnProGlnProSerThrGlySerLeuThrGlySerProHisLeuSerGluLeuSe 484
Db 139 -GGGCCCGCGAGCCCGCGGAGGCGCGGTTCATCCACAGCGTCCTGCACCATGCC 197
Qy 484 rValAsnSerGlnGlyValAlaProAlaAsnValThrLeuSerProAsnLeuSerPr 504
Db 198 CTCAGACTCAGAAAGACAGCAGCTCCCTCAGCAGTGT----- 234
Qy 504 AsnPheTrpLysGlnLysSerProLeuIleSerProLeuLeuAsnArgAlaCysProAr 524
Db 235 -GGCATTCAGGGAAGGCGCGTCCCAACACCCCTCTCACTGACGAAAGTAAATGA 293
Qy 524 g-----ThrAspArgLysArgGluArg 532
Db 294 GAAGGTGAGAAACCTTCATTCAGCTGCGGTGATATACCGGGAGAGAAAGAGCTCG 353
Qy 532 gArgLys--ArgPheProThrAspLysAlaTyrPheIleAlaLysGluValSerThrTh 551
Db 354 CAAGCGCTGCGCTTTGCTACGATGCGACGCGCTTTGACAAAGAGGCCCTTACACAGGCT 413
Qy 551 rGluArgThrTyr-----LeuLysAspLeuGluVal 561
Db 414 GAATCTGATATATAGAGGCTGAAGATTCAGTGGGTGGAGCCATGTGCACTGGAGAG 473
Qy 561 IleThrSerTrpPheGlnSerThr-----ValSerLysGluAspAl 575
Db 474 CTTGCAAGAACCAAGACACCAACGCTTGAAGAAGGTGAGAGAGGTGGCCAGAGAGCTGA 533
Qy 575 AsnProGluAlaLeuLysSerLeuIlePhe-----ProAsnPh 588
Db 534 CTTTACACACACTCCACACAGCCGCTCTGATGACACAGACTCGGCTGAAGATGACGT 593
Qy 588 eGluProLeuHisLysPheHisThrAsnPheLeuLysGluIleGluGlnArgLeuAlaLe 608
Db 594 GGAACATGCTGAGGGGAGAAATGGGACAGCTCTCGGAGGAAACCTGTCGACAGAGTC 653
Qy 608 uTrpGluGlyArgSerAsnAlaGlnIleArgAspTyrGlnArgIleGlyAspValMetLe 628
Db 654 ATGGAGAGCATGAAAGCGCTCCACAGAGAGACAAAGAGATCGGTGACCTCCGTC 713
Qy 628 uLysAsnIleGlnGlyMetLysLysLeuAlaAlaHisLeuTrpLysHisSerGluAlaLe 648
Db 714 CCAGCAGCAGCAGGTGTGAAGCAACAGGCTCA-----TCCGAGATTCT 758

```

```

RESULT 11
US-10-559-415-189
; Sequence 189, Application US/10559415
; Publication No. US20060100132n1
; GENERAL INFORMATION:
; APPLICANT: Astrazeneca AB et al
; TITLE OF INVENTION: Diagnostic Method
; FILE REFERENCE: 101073-1P WO
; CURRENT APPLICATION NUMBER: US/10/559,415
; PRIOR FILING DATE: 2005-12-06
; PRIOR APPLICATION NUMBER: 0313081.2
; NUMBER OF SEQ ID NOS: 191

```



```

Qy 648 uGUlAlaUeUgUaNgUlylUeUySeSeSeArGrUeUgUaNPheCyArGrPh 668
|         |||         :|:         :|:         :|:         :|:
Db 759 CAACAACTGTATAC-----ACGGCCATGACAGATTGAGGTGGTCAAGAGACTA 812
Qy 668 eGUlUeUgUaUyValCyUyTyUeUProUeUaNPheUeUaGrProUeUaUy 688
|         |||         :|:         :|:         :|:         :|:
Db 813 TGAGCCCTTCGGAAGGTACAGTGAAGAAAGTCCGATCCAGATGAGACTGAGCCG 872
Qy 688 gUeUeUeUyUyUyUyUyUyUyUyUyUyUyUyUyUyUyUyUyUyUyUyUy 708
|         |||         :|:         :|:         :|:         :|:
Db 873 CTTGAGACAGCTGGGGGAGGAAACAGCGGTCTGTAAGCAG----- 915
Qy 708 aAsPheArGrArGrArGrAlaAlaUeUaUgUuUleUThGrUeUeUaUaUgUeU 728
|         |||         :|:         :|:         :|:         :|:
Db 916 -----ACAGAGATGCTGACCCAGCGAG 938
Qy 728 sGUlUeUeUyUyUyUyUyUyUyUyUyUyUyUyUyUyUyUyUyUyUyUyUy 740
|         |||         :|:         :|:         :|:         :|:
Db 939 GGAACAGCGCCATCCAGCTGACAGCAAGTCCCTCTCCCTGAGAGGTTTGAAGCGAT 998
Qy 740 u---HieGUeUeUyUyUyUyUyUyUyUyUyUyUyUyUyUyUyUyUyUyUy 759
|         |||         :|:         :|:         :|:         :|:
Db 999 CCACCATGAGCTGAACAG----- 1017
Qy 759 uPheUleArGrUeUyUyUyUyUyUyUyUyUyUyUyUyUyUyUyUyUyUyUy 779
|         |||         :|:         :|:         :|:         :|:
Db 1018 -----GCCACGGCGCAGAACAGAGACTGACGAGGATGGA 1055
Qy 779 ePheUeUeUeUeUeUeUeUeUeUeUeUeUeUeUeUeUeUeUeUeUeUeUe 798
|         |||         :|:         :|:         :|:         :|:
Db 1056 GCTGCTCAGTCAAGAGCTGACCGAGCTGAGAACCCAGCGTGAAGACAGCA----- 1107
Qy 798 nPheUyUaUHiGUyUeUProUeUyUyUyUeUThUeUeUyUyUyUyUyUyUyUy 818
|         |||         :|:         :|:         :|:         :|:
Db 1108 -----AAGAGTCCGAG----- 1119
Qy 818 uTPRGlYValProHieCyUeUThUeUArGrUyUyUyUyUyUyUyUyUyUyUy 835
|         |||         :|:         :|:         :|:         :|:
Db 1120 -----AAATACAGGGAGAGAGCGGAGCGCTGTACAGCGATA 1157
Qy 836 -----ValAlaAlaSeSeSeArGrUeUeUyUyUyUyUyUyUyUyUyUyUy 854
|         |||         :|:         :|:         :|:         :|:
Db 1158 CAAGCTCATCATGAGTGAAGCGTGAACAGCTATCTGAGCTGACAGCTGACAGCGA 1217
Qy 854 aUeUeUeUeUeUeUeUeUeUeUeUeUeUeUeUeUeUeUeUeUeUeUeUe 874
|         |||         :|:         :|:         :|:         :|:
Db 1218 AGTGAAGCTGGCCAGATCCAG-----CTCAAGACAGCACATC 1256
Qy 874 cAsPAsUyUySeSeSeProArGrUaUaUaUaUaUaUaUaUaUaUaUaUaUa 892
|         |||         :|:         :|:         :|:         :|:
Db 1257 TGAGAAAGAGCGGCGCAATGAGATGAGAGCGCTCGGAGATCAAGACAGCGTAC 1316
Qy 892 pUeUSeUeUaUSeUyUyUyUyUyUyUyUyUyUyUyUyUyUyUyUyUyUyUy 901
|         |||         :|:         :|:         :|:         :|:
Db 1317 AATGATGCTGGAGAGCCAAACAGAG 1344

```

RESULT 12

```

US-10-505-928-450
; Sequence 450, Application US/10505928
; Publication No. US20060088532A1
; GENERAL INFORMATION:
; APPLICANT: Ludwig Institute for Cancer Research et al.
; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
; FILE REFERENCE: 28967/39178
; CURRENT APPLICATION NUMBER: US/10/505,928
; PRIORITY FILING DATE: 2004-08-27
; PRIOR APPLICATION NUMBER: US 60/363,019
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: PatentIn 3.2
; SEQ ID NO 450
; LENGTH: 11185
; TYPE: DNA

```

```

; ORGANISM: Homo sapiens
US-10-505-928-450
Alignment Scores:
Pred. No.: 4.28
Score: 137.50
Percent Similarity: 33.1%
Best Local Similarity: 19.5%
Query Match: 2.5%
DB: 6
Gaps: 50
US-09-555-342B-2 (1-1045) x US-10-505-928-450 (1-11185)
Qy 51 GInGUlUaUeUgUaUyUaUProGInArGrAlaProGly-----UySeUaUeUaUeUaUa 68
|         |||         :|:         :|:         :|:         :|:
Db 6633 GAAAGATGATTTGAATCCCTCABAACTCCTCGCAAGACAGAACAAACATCTTGATTC 6692
Qy 69 --ValCyUaUaUyUyUyUyUyUyUyUyUyUyUyUyUyUyUyUyUyUyUy 83
|         |||         :|:         :|:         :|:         :|:
Db 6693 CAGACATTTACTGAACCTGAACCAAAACAGATTTATCTGTACTTACCAAGAA 6752
Qy 84 GlUeUeUeUeUeUeUeUeUeUeUeUeUeUeUeUeUeUeUeUeUeUeUeUe 99
|         |||         :|:         :|:         :|:         :|:
Db 6753 ACTTACAGTGTATGAATGAAGAAAGAGAGACACTTCTTGTAGTACATGCTACT 6812
Qy 100 Pro-----UeUaUyUyUyUyUyUyUyUyUyUyUyUyUyUyUyUyUyUyUy 108
|         |||         :|:         :|:         :|:         :|:
Db 6813 CCAGATCCAGATCCAAATGCTTGAATCTTACACAACTCCCTGAGCTAGCAAAAG 6872
Qy 109 UyUeUeUeUeUeUeUeUeUeUeUeUeUeUeUeUeUeUeUeUeUeUeUeUe 126
|         |||         :|:         :|:         :|:         :|:
Db 6873 TCACATTTTCTTACTACTGCTATTAAGTGAATCTTACACAGTGAATGATGTC 6932
Qy 127 UeUeUeUeUeUeUeUeUeUeUeUeUeUeUeUeUeUeUeUeUeUeUeUe 141
|         |||         :|:         :|:         :|:         :|:
Db 6933 ACAGATTCACCAATCAAAAGAAAGTACAAACAT----- 6971
Qy 142 GInAsPheUaUgUyUyUyUyUyUyUyUyUyUyUyUyUyUyUyUyUyUyUy 160
|         |||         :|:         :|:         :|:         :|:
Db 6972 -----TTTCCGAAAGCATGAGACCAACATTCAGAGCTAGTACGCTTATTC 7025
Qy 161 SeUeUeUeUeUeUeUeUeUeUeUeUeUeUeUeUeUeUeUeUeUeUeUeUe 171
|         |||         :|:         :|:         :|:         :|:
Db 7026 TCTGAGCTGGATCAGAGAGAAAGTTTTACCTACTCTCAACAGAGCTAGGAATTTT 7085
Qy 172 AsPGUaUaUeUaUyUyUyUyUyUyUyUyUyUyUyUyUyUyUyUyUyUyUy 191
|         |||         :|:         :|:         :|:         :|:
Db 7086 ACTGAAGTG-----GAACAAATCAATTAACATTTATCCACACTTCA 7133
Qy 192 UeUgUeUeUeUeUeUeUeUeUeUeUeUeUeUeUeUeUeUeUeUeUeUeUe 202
|         |||         :|:         :|:         :|:         :|:
Db 7134 GTGGAAGATACCTCAAGTGAACAAATTTGAAGACTTTTAAACAGATGGAATGGCAAA 7193
Qy 203 HieUeUyUyUyUyUyUyUyUyUyUyUyUyUyUyUyUyUyUyUyUyUyUy 213
|         |||         :|:         :|:         :|:         :|:
Db 7194 GAGTTGACACCACTGATCTCAAAACAGATCTTGAAGTGTAGTGGGTCAAGCAGC 7253
Qy 214 -----GInUeUeUgUuUeUaUaUaUaUaUaUaUaUaUaUaUaUaUa 232
|         |||         :|:         :|:         :|:         :|:
Db 7254 ACAACATTAATTAATTT-----TTAAGTGAACAGTGA----- 7286
Qy 233 UyAsPArGrUyUyUyUyUyUyUyUyUyUyUyUyUyUyUyUyUyUyUyUyUy 252
|         |||         :|:         :|:         :|:         :|:
Db 7287 -----GAGAAAGAGACCAAGCGTGAACCTCTTCTCCAGGAGATC----- 7331
Qy 253 GInGUyUyUyUyUyUyUyUyUyUyUyUyUyUyUyUyUyUyUyUyUyUyUy 272
|         |||         :|:         :|:         :|:         :|:
Db 7332 ---GAGATCTCAAAATCAGACTGTCAAGTGGCGAGAA-----GAAATCCAGACTAGT 7382
Qy 273 ArGUyArGrPheUeUeUyUyUyUyUyUyUyUyUyUyUyUyUyUyUyUyUyUyUy 292
|         |||         :|:         :|:         :|:         :|:
Db 7383 AGACCAACAACTAATCTGAACAGACTCTTACAGAAATTTCTTCAACAGCAAGAAATTAAC 7442

```

```

Qy 293 GluPheUmeMetAlaSerArgAspPheCysLysSerPheTrpValIleCysValGluHis 312
Db 7443 GAAACAAACACCTCTACTGATTTTCTGGCTAGACTTATGCTTT----- 7490
Qy 313 HisAlaPhePheArgLeuPheGluGluProLysProLysProValLeuPhe--- 331
Db 7491 ---GAATGGCCAAAGATTTGTTATCATCAGACCCAAACCATCTGACTGTATATGAA 7547
Qy 332 ---SerArgLysSer----- 342
Db 7548 CTTTCTGGAGAAGATCTGGAGAACTGATATTGTGATTCATTTCACTTCCATCTGCAACT 7607
Qy 343 Thr-----GlnLysGlnValLeuAspTrpValLysGluGlyHisLysLys 358
Db 7608 ACTCAGGCAACACAGAAAGAAAGACGACCCACATTTGTTCTGATGGGCTCCCTGAA--- 7664
Qy 359 ValGlnPheGluArgLysHisSerLysLysHisSerLysLeuArgSerLeuAlaSerGln--- 377
Db 7665 ---AATCATCTGAGGTCGCAAGCCGCTAAAGCTTAAAGCTTACTGCTGATGGA 7709
Qy 378 ---ProThr-----GluLeuAsnSerGluValLeuGluGlnSerGln 391
Db 7710 TTCCCAACAGTTTACAGTATGCTGCTCTCTCATTCAGACGAAACMAAGCTCCCTGAT 7769
Qy 392 SerThrSerLeuThrPheGlyGluGlyAla-----GluSerProGlyGlyGlnSerCys 409
Db 7770 CCAACTAGCACACCTGCAATAATACAGTGCATATGAGAGCTCCACAGACGTAAGTTCCAA 7829
Qy 410 ArgArgLysGluProLysValSerAlaGlyGluProGlySerHisProSerProAla 429
Db 7830 GACGTTTACAGGAATTCAGAGATTCACCTTAAACCTAAC----- 7871
Qy 430 ProArgArgSerProAlaGlyLysLysGlnAlaAspGlyAlaAlaSerAlaProThrGlu 449
Db 7872 ---AGAAAAAAACCCACTGAATAATTTATCATAGAC-----CTGGACAAA 7913
Qy 450 GluGluGluGluValValLysAspArgThrGlnGlnSerLysProGlnProProGlnPro 469
Db 7914 GAGACAAAGATTTAATATTCG----- 7934
Qy 470 SerThrGlySerLeuThrGlySerProHisLeuSerGluLeuSerValAsnSerGlnGly 489
Db 7935 ---ACAATTACAGAGATGATCATCTTGAATTTCTA----- 7967
Qy 490 GlyValAlaProAlaAsnValThrLeuSerProAsnLeuSerProAspThrLysGlnAla 509
Db 7968 ---CCTGACCTGACATCGGATAAATACTATCATATGATATTCATCATCT 8015
Qy 510 SerProLeuLeuSerProLeuLeuAsnAspGlnAlaCysProAspThr----- 525
Db 8016 AATCCTGTATGAAGACATCTTGGAAATGCMAACAGATATGATACAGAGTACCATCA 8075
Qy 526 ---AAspArgLysAspGlyGluLysArgLysArgPheProThrAspLysLeuLys 542
Db 8076 GAACACATGACAGTAATGATGAAGTAATGATGACACACTCAAGTCAAGAGATCAT 8135
Qy 543 PheIleAlaLysGluValSerThrThrGluArgThrTyLysLeuLysAspLeuGluValIle 562
Db 8136 GAGCAGCCTGTCAACCTTTCTTAATCTGAGAAACATTTGAGGGCTGCTGATGTTCTG 8195
Qy 563 ThrSerTrpPheGlnSerThr-----ValSerLysGluAspAlaMetProLys 578
Db 8196 GCTAGCTACACTACAGCAACATGATATCATAGACTTTATGAAGATAG-AAAGCCAACT 8254
Qy 579 AlaLeuLysSerLeuLeuPheProAsnPheGluProLysHisLysPheHisThrAspPhe 598
Db 8255 AGATCAATGGGCTTCACTTCAACAATGGGATCCCTGC---TCTTACACAGAAACAGA 8311
Qy 599 LeuLysGluLysGluGlnArgLeuAlaLeuTrpGluGlyArgSerAsnAlaGlnIle-Arg 618
Db 8312 ATTAGACCTTTT-----ACTTCCACAGGCAACATCCCT 8344
Qy 618 GAspTrpGlnArgLysGluAspValMetLeuLysAsnIleGlnGlyMetLysHisLeuAl 638

```

```

Db 8345 GCCAATTCCTCGTAAGTCTGCCACAGTATTTCCAGAGTTGTAAGAAATAAA----- 8396
Qy 638 AlaHisLeuTrpLysHisSerGluAlaLeuGluAlaLeuGlnLysSerSer 658
Db 8397 ---GCTGAAGCA---AAAGCCCTGATACATGTTTGAATCAAG 8434
Qy 658 ArgArgLeuGluAsnPheCysArgAspPheGluGlnLysValCysTyLysProLe 678
Db 8435 CACTTGTCTGATGCT----- 8450
Qy 678 ValAsnThrPheLeuLeuArgProLeuHisArgLeuMetHisTyLysGlnValLeuGluArg 698
Db 8450 ----- 8450
Qy 698 GluCysLysHisHisProProSerHisAlaAspPheArgAspCysArgAlaAlaLeuAl 718
Db 8451 -----CAAGCTATTGC 8461
Qy 718 ArgLysLeuThrGluMetValAlaGlnLeuHisGlyThrMetLysMetGluAsnPheG 738
Db 8462 AGACCAAGTAAGTAATATACCA-----ACATGGGCCAATTTGAAAGACTCA 8509
Qy 738 nLysLeuHisGlyLeuLysLysAspLeu-----IleG 749
Db 8510 GGAGAGATGATGAAGACAAABAAACATGCTGCTCTTTTCAAGCAAGATTTCTTTCAGG 8569
Qy 749 YLysAspAsnLeuValAlaProGlyArgGluPheIleArgLeuGlySerLeuSerLysLe 769
Db 8570 AGCTAGAGAGGACTAGTACATACCTATCCCTATCTAAGTTCTCTATCCACAC----- 8624
Qy 769 uSerGlyLysGlyLeuGlnGlnArgMetPhePheLeuPheAsnAspValLeu----- 786
Db 8625 ---CTTATGATCAGAGTGAACAGAGCTGCTATGATGATGAAGATC 8671
Qy 787 ---LeuTyThrSerArgGlyLeuThrAlaSerAsnGlnPheLysValHisG 803
Db 8672 CAATCCCCCATATTAACATGATACATTAAGACTTTCACATTTGGAGAGTGTCTTC 8731
Qy 803 YGlnLeuProLeuTrpGlyMetThrIleGluGlnSerGluAspGluTrpGly----- 820
Db 8732 TCAGACACATCATCTCCCTCATCTACTCAGAGAGTGAAACCTCTGACACACAGA 8791
Qy 821 -ValProHisCysLeuThrLeuAspGlyGlnArgGlnSerIleIleValAlaAlaSerSe 840
Db 8792 GATCCCCCAGCCAGCTGCTGCCAGGA-----ATGAACGTGGCTATCTGT 8839
Qy 840 ArgSerGluMetGluLysTrpValGluAspIleGlnMetAlaIleAspLeuAlaGluLys 860
Db 8840 AATGTCACACAGATTTCTTT---AAGAAATTCATGTAATTTGAAGCAACTTCA 8896
Qy 860 sSerSerSerProAlaProGluPheLeuAlaSerSerProAspAsnLysSerProAs 880
Db 8897 ACCATCAAGT-----GAGGAATACCTTCACATAGAGAGCTCCCTTATATCTCTGA 8950
Qy 880 PGLuAlaThrAlaAlaAspGlnGlnLysSerGluAspAsp----- 892
Db 8951 CACAAAATTA-----GAACCTCAGAAAGATGAAGTAAACCTGAGTTATTAAGAA 9001
Qy 893 -LeuSerAlaSerArgThrSerLeuGluArgGlnAlaProHisArgGluAsnThrMetVal 912
Db 9002 AATGGAAGCTTCTCCACAGAACTT----- 9026
Qy 912 HisValCysTrpHisArgAsnThrSerValSerMetValAspPheSerIleAlaValG 932
Db 9027 -----ATTGCTGTGGA 9037
Qy 932 ValAsnGlnLeuSerGlyAsnLeuLeuArgLysPheLysAsnSerAsnGlyTrpGlnLysLe 952
Db 9038 AGGAACCT-----GAGATTTCTCAAGATTTCCAAACAA-AAACGATGGTCAAGTTT 9087
Qy 952 uTrpValValPheThrAsnPheCysLeuPhePheTyLysSerHisGlnAspAsnHisAsp 972

```

Db 9088 CTGG-----AGAACATCAAGATG--TTTCC 9112
 Qy 972 oleuAlaSerIeuProIeuLeuGlyTyrSerLeuThrIleProSerGluSerGlu 990
 Db 9113 CACCATTAACACCTGAGCTGGACCTGTTATTACACTCCCATGAAATTGAA 9167

RESULT 13

US-11-118-524-1
 Sequence 1, Application US/11118524
 Publication No. US20060088647A1
 GENERAL INFORMATION:
 APPLICANT: GU, WEI
 TITLE OF INVENTION: ARF-BP1 AS MEDIATOR OF P53-DEPENDENT AND INDEPENDENT TUMOR
 FILE REFERENCE: 19240-497US2
 CURRENT APPLICATION NUMBER: US/11/118,524
 CURRENT FILING DATE: 2005-04-29
 PRIOR APPLICATION NUMBER: 60/610,506
 PRIOR FILING DATE: 2004-09-15
 NUMBER OF SEQ ID NOS: 22
 SOFTWARE: PatentIn version 3.3
 SEQ ID NO 1
 LENGTH: 13128
 TYPE: DNA
 ORGANISM: homo sapiens
 FEATURE:
 NAME/KEY: misc feature
 LOCATION: (3370)..(3372)
 OTHER INFORMATION: CODES FOR A PROLINE AS COMPARED TO A LEUCINE AT AMINO ACID
 OTHER INFORMATION: RESIDUE 1124 OF SEQ ID NO:2
 FEATURE:
 NAME/KEY: misc feature
 LOCATION: (7918)..(7920)
 OTHER INFORMATION: CODES FOR A LYSINE AS COMPARED TO A THREONINE AT AMINO ACID
 OTHER INFORMATION: RESIDUE 2640 OF SEQ ID NO:2
 FEATURE:
 NAME/KEY: misc feature
 LOCATION: (7921)..(7923)
 OTHER INFORMATION: CODES FOR A PHENYLALANINE AS COMPARED TO A LEUCINE AT AMINO ACID
 OTHER INFORMATION: RESIDUE 2641 OF SEQ ID NO:2
 FEATURE:
 NAME/KEY: misc feature
 LOCATION: (7924)..(7926)
 OTHER INFORMATION: CODES FOR A VALINE AS COMPARED TO A SERINE AT AMINO ACID RESIDUE
 OTHER INFORMATION: 2642 OF SEQ ID NO:2
 FEATURE:
 NAME/KEY: misc feature
 LOCATION: (7927)..(7929)
 OTHER INFORMATION: CODES FOR A Cysteine AS COMPARED TO A Serine AT AMINO ACID
 OTHER INFORMATION: RESIDUE 2643 OF SEQ ID NO:2
 FEATURE:
 NAME/KEY: misc feature
 LOCATION: (9955)..(9957)
 OTHER INFORMATION: CODES FOR AN ASPARTIC ACID AS COMPARED TO A GLYCINE AT AMINO ACID
 OTHER INFORMATION: RESIDUE 3319 OF SEQ ID NO:2
 FEATURE:
 NAME/KEY: misc feature
 LOCATION: (11293)..(11295)
 OTHER INFORMATION: CODES FOR AN ALANINE AS COMPARED TO A VALINE AT AMINO ACID
 OTHER INFORMATION: RESIDUE 3765 OF SEQ ID NO:2
 US-11-118-524-1

Alignment Scores:

Pred. No.: 5.72 Length: 13128
 Score: 137.00 Matches: 199
 Percent Similarity: 31.9% Conservative: 110
 Best Local Similarity: 20.5% Mismatches: 326
 Query Match: 2.5% Indels: 336
 DB: 7 Gaps: 47

US-09-555-342B-2 (1-1045) x US-11-118-524-1 (1-13128)

Qy 121 ProProAspHisThrGlnLeuGlnGluLeuThrArgTyrLeuPheAlaLeuGlnVal 140

Db 9076 CTTCCA-----GCCATTGAGAGAGAGTACTGGCACAGCAGAGCTGAGCACAG 9126
 Qy 141 LyeGlnAspLeuAlaGlnGlyArgLeuThrCysAspAspThrSer----- 155
 Db 9127 CGACGAGAACTTACAGCAAGAT-----GCCAGCTCACACCCCTTATGACCTGTGACC 9180
 Qy 156 ---AlaAlaLeuLeuIleSerHisIleValGlnSerGluIleGlyAspPheAspGlu--- 173
 Db 9181 TTGATCCAGACTGTGGCCCTGAGACCTGGCCGCGTAGTCTCTTAGAGGATATGAGGACAGT 9240
 Qy 174 -----AlaLeuAspArgGluHisIleValAlaLeuAspGlu 184
 Db 9241 GTGTGACTGTGATGCCACCTGCATTCAGCTGACGAGCTCAAGCCCTGACAGCAGACCAA 9300
 Qy 185 TyrIleProGlnIleAspAlaLeuGluAspArgIleValGluPheHisIleAspHis--- 203
 Db 9301 GAAGCCCGGACGACGACCTCATGACATGACGCTGTG-----TTTGGGACACATGACACC 9354
 Qy 204 -----IleGlyGlnThrProAlaGluSer----- 211
 Db 9355 TCCGCACTCTGTGCTATTCTCCGAGCCGCTTTCACACGCTTAAGTGCACACCT 9414
 Qy 212 -----AspPheGlnLeuLeuGluIleAlaArg-----ArgLeuGluMetTyrGlyIle 227
 Db 9415 GGGGTCCAGTATATCGCTTGTCTGTGACAGAGTGCACACTTCCAGATGGGGGTATGC 9474
 Qy 228 ArgLeuHis-----ProAlaValAspArgGluGlyThrTyrIleAspLeu----- 242
 Db 9475 AGCAGCATTAACAGGCTTGTGACGATGATGATATCTCTCGCCCTCCAGAGACGG 9534
 Qy 243 -----AlaValAlaAspThrGlyIleLeuValPheGlnGlyPheThr 256
 Db 9535 CTCCTTGTGACACAGAGAGCCCTTCTGTCTGTGTGCTTACTTTTGTGATGAGCCA 9594
 Qy 257 LyeIleAspAlaPheAspThrAlaValAlaVal-----ArgLysLeuSerPhe-----LysArg 273
 Db 9595 AACTCAATCTAGCCGCTTACACCCAGTACTAGAAATCTGTCTACATGCCACAGACC 9654
 Qy 274 LysArgPheLeuIleLysLysLeuArgProAspAlaAspSerLysArgGlnAspThrLeu 293
 Db 9655 GGCACCTGGGTCACTCCG-----AGTGTGCTTCC 9684
 Qy 294 PheLeuMetAlaSerArgAspPheCysLysSerPheThrLysIleCysValGlnHis 313
 Db 9685 ATCTTGAGCGCAGCAGTAGAGT-----GAGCTATGCAATGAAACACC 9729
 Qy 314 AlaPhePheArgLeuPheGluGluProLysPheLysProLysProValLeuPheSerArg 333
 Db 9730 AAACCTCACTCAAGTAGGAGAAAGGCGCAAAAGTCGAGCAAG-----AGCTGT 9777
 Qy 334 GlySerSer-----PheArgPheSerGlyArgThr 343
 Db 9778 GGGTCAGAGCCATGAGAAACGTCCTCCCTGACCTGTACACAAGATGAGTCAAAAGAGC 9837
 Qy 344 GlnLysGln----- 346
 Db 9838 TCCACACAGCTTCTCGGCTCTCATGATCCATGATGACAGCCCTAGGCTGACGAGCTAAT 9897
 Qy 347 ValLeuAspTyrValLysGluGlyGlnHisLysLysLeuValGlnPheGluArgHis--- 365
 Db 9898 ATATTTCAGATCCAGGCTTCCAGGGGCGGTAAACATACCGAG-----AAGCATGCA 9948
 Qy 366 -----SerLysIleHisSerIleArgSerLeuAlaSerGlnProThrGluLeuAsn 382
 Db 9949 AGCGGTGACTCCACCGTCCATCATCCCAAGCTGCACCTGTGTCTGACAGACGCTT 10008
 Qy 383 SerGlnValLeuGluGln-----SerGlnGlnSerThr 393
 Db 10009 TTGGATATCACTCATTTGATGGCCAGGATATTTCCAGCCACTTCCACAGCAGCGGAGCC 10068
 Qy 394 SerLeuThrPheGlyGluLysAlaGluSerProGlyGlyGlnSerCys----- 409

Db 10069 AAGAAACAACCTGTGAGTGTATGCGGAAAGGCAATAAGGCTGTAGCCATGCTCC 10128
 Qy 409 ----- 409
 Db 10129 TCACAGTCTCCAGCAGTGGCATTTGCACAGACTTCTGGGACTTATGGTAAACTGGAC 10188
 Qy 410 -----ArgArgGlyLeuPro-----LysValSerAlaGly 420
 Db 10189 AACATGATGTCCACCGGAAAGGCAAGAACTCCGTAAGTGCAGTCCAGTACGCGCTGGC 10248
 Qy 421 GluProGly----- 423
 Db 10249 GGTAGGGGGAACCTCTCCATACAGCTCGAGGCTCTCCACTGGGGCAGCTCATGAAC 10308
 Qy 424 -----SerHisPro----- 426
 Db 10309 ATGTGTTCACACCCAGTCATCCGCGGAGCTCTCTTAACTGAGAACTCCAGACTC 10368
 Qy 427 -----SerProAlaProArgSerProAlaGlyAsnLysGln 439
 Db 10369 CTTTCTCATCTCAATGCTCTCCAGAAACAGGTGTCAAGACAGGCTAATCT 10428
 Qy 440 AlaAspGlyAlaAlaSerAlaProThrGluGluGluValValLysAspArgThr 459
 Db 10429 GGCAGCGGTGCTCTCCACACCACTGCACCTCAACCATCATCACCAACCACT 10488
 Qy 460 GlnGlnSer---LysProGlnProProGlnProSerThrGlySerLeuThrGlySerPro 478
 Db 10489 GCGCGCTTCACACGACCCACACCCCTCATGACCCACC---CCTGTCACTTCTGCTCA 10545
 Qy 479 HisLeuSerGluLeuSerValAsnSerGlnGlyValAla----- 492
 Db 10546 GCGCTGTGCTGCCACGCGCATTTCCACCATGTCTGAGTCTTGCACACCACTGACT 10605
 Qy 493 ---ProAlaAsnValThrLeuSerProAsnLeuSerProAspThrLysGlnAlaSerPro 511
 Db 10606 ACCCCACGACTGCTACCTACCTATGTTCAATTTCTCCACTACTACGAGCAGCAA--- 10662
 Qy 512 LeuLeuSerProLeu---LeuAsnAspGlnAlaCysProArgThrAspAspGluAspGlu 530
 Db 10663 ---TCTCCAGCAAGGTAGTGAATGGGCGACGACAGCTACAGAC--- 10704
 Qy 531 GlyArgArgLysArgPheProThrAspLysAlaLysPheLeuAlaLysGluValSerThr 550
 Db 10705 -----TTTAAGATGGTGTCTCT 10722
 Qy 551 ---ThrGluArgThrTyLeuLysAspLeuGluValLethrSerTrpPheGlnSer 568
 Db 10723 GGCCTCACTGAAACCAAGCTACAGCTCTGTAGAGGTGTGACATCCACTCTGTGTTCT 10782
 Qy 569 ThrValSerLysGluAspAlaMetProGluAlaLeuLysSerLeuLethrProAsnPhe 588
 Db 10783 GAGGAGGCTTAAAGGATGCA----- 10803
 Qy 589 GluProLeuHisLysPheHisThrAsnPheLeuLysGluLeuGlnArgLeuAlaLeu 608
 Db 10804 -----GCCAAGCTACTACTGCAGCTC----- 10824
 Qy 609 TrpGluGlyArgSerAsnAlaGlnLLeaArgPheTyGlnArgLLeuLysAspValMetLeu 628
 Db 10825 ---TCCGGGGGAGCTCGGAGCCGGAGC-----ACTGTTCTCAAG 10863
 Qy 629 LysAsnLLeuGlnGlyMetLysHisLeuAlaAlaHisLLeuTrpLysHisSerGluAlaLeu 648
 Db 10864 CTGCTACTGATGAGCCCGCATCTGGTTATACCTTTGTAAACAAATAGTACCTCG 10923
 Qy 649 ---GluAlaLeuGluAsnGlyLLeuLysSerSerArgArgLeuGluAsnPheCysArg 666
 Db 10924 CTGGCCAGCTCGCGGAATACAACTCGACAGACAGCGGCGGCCAA----- 10971
 Qy 667 AspPheGluLeuGlnLysValCysTyLeu---ProLeuAsnThrPheLeuLeuArgProLe 686
 Db 10972 ---TGTGAACCTCTCTCTGATGCGCTGAGTGAAGACAG 11010

Qy 686 WHisArgLeuMetHisTyTrLysGlnValLeuGluArgLeuCysLysHisHisProProse 706
 Db 11011 CCACAGA-----CCACAGA 11025
 Qy 706 rHisAlaAspPheArgAspCysArgAlaAlaLeuAlaGluLethrGlnMetValAlaGly 726
 Db 11026 CTGAAG---GCCAAATGTCAGACAGCTTTG----- 11053
 Qy 726 nLeuHisGlyThrMetLLeuMetGluAsnPheGlnLysLeuHisGluLeuLysLysAs 746
 Db 11054 ---ACATGGCTGAGAAATG---TGGTAATTTGGGCACTCAAGACGGA 11094
 Qy 746 nLeuLLeuGlyThrLeuAsnLeuValProGlyArgGluProLethrLeuArgLysSerLe 766
 Db 11095 CTTTG---GGT-----GCCCGGAG---CTCAGCTGCTCTTAT 11129
 Qy 766 userLysLeuSerSerLysGlyLeuGlnArgMetPhePheLeuPheAsnAspValLe 786
 Db 11130 GTCCATGTGACATCCAAAGCATCTACCCAGAACTTCTTGAAGGTACTACAGTCAT 11189
 Qy 786 uLeuTyT-----ThrSerArgGlyLeuThrAlaSerAsnGlnPheLysValHis 802
 Db 11190 CATCCAGCTCCGGAGCAGACAGCGCCGG-----GCTAAACAAGAAAGCCAGACAGAC 11240
 Qy 802 sGlyGlnLeuProLeuTyTrGlyMet-----ThrLLeuGln 814
 Db 11241 AGGCAAGCTAGTGTCTCCGGTTTAAAGCTCAGCTACGACATCCAGCAGCTGCTCGCA 11300
 Qy 814 userGluAspGluTrpLysValProHisCysLeuThrLeuArgGlyGlnArg----- 831
 Db 11301 GCTGAGGCTGAGGCTGATATCCATTAACAATGTACAGAGGTCAAAAGCGCGGAG 11360
 Qy 832 ---GlnSerLLeuValAlaAlaSerSerArgSerGlnMetGluLysTrpValGluAs 850
 Db 11361 ACAGCAACAAGCAGCAAGCTCGAGTCTACGCTCAGAGCGCTCTCCGAGAGAGA 11420
 Qy 850 pLleGlnMetAlaLLeuPheLysGluLysSerSerSerProAlaProGluPheLeuAl 870
 Db 11421 ATCACCCATGATGTGACCAAGCCATCTCCAGTGTCAAGATCACTCAATCCATTCCTC 11480
 Qy 870 aserSerProPro-----AspAsnLysSerProAspGluAlaThrAl 884
 Db 11481 CGATGAAACCCACAGGCGGAGAAAGAAAGAAAGAAAGCAACCTAGATTACCCCTGCT 11540
 Qy 884 alaAspGlnGlnLysSerGluAspAspLeu 893
 Db 11541 CAGCGACAGCTGAGTTTGACGAGCTG 11568
 RESULT 14
 US-11-217-529-77055
 ; Sequence 77055, Application US/11217529
 ; Publication No. US20060099612A1
 ; GENERAL INFORMATION:
 ; APPLICANT: SUNTORY LIMITED
 ; APPLICANT: NAKARO, YOSHIHIRO
 ; APPLICANT: NAKAMURA, NORIHISA
 ; APPLICANT: KODAMA, YUKIKO
 ; APPLICANT: FUJIMURA, TOMOKO
 ; APPLICANT: ASHIKARI, TOSHIHIKO
 ; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
 ; FILE REFERENCE: S-38-285
 ; CURRENT APPLICATION NUMBER: US/11/217,529
 ; PRIOR FILING DATE: 2005-09-02
 ; PRIOR APPLICATION NUMBER: US 10/932,182
 ; NUMBER OF SEQ ID NOS: 197023
 ; SOFTWARE: PatentIn version 3.3
 ; SEQ ID NO 77055
 ; LENGTH: 5040
 ; TYPE: DNA
 ; ORGANISM: Saccharomyces pastorianus
 ; US-11-217-529-77055

Alignment Scores:

Pred. No.:	1.75	Length:	5040
Score:	136.50	Matches:	180
Percent Similarity:	35.1%	Conservative:	145
Best Local Similarity:	19.5%	Mismatches:	366
Query Match:	2.5%	Indels:	235
DB:	7	Gaps:	41

US-09-555-342B-2 (1-1045) x US-11-217-529-77055 (1-5040)

```

Qy 37 GlyLeuValSer-----IleValIleIleMetLeuSerPheThrGlnGlu 52
    |||||
Db 1081 GGTAACTATTCTGTGATATCAAGTCTGTAAAGACACTGATTAAGGAAGAAATCA 1140
    |||||

Qy 53 AlaPheGluValProGlnArgAlaProGlyValValLeuLeuSerAlaValCysAsnHis 72
    |||||
Db 1141 AAATTTCATTTACGAATTCAAATG----- 1164

Qy 73 LeuAsnLeuValGluGluAspTyrPheGlyLeuGluPheProAspHisLys-----Lys 90
    |||||
Db 1165 -----GAGGATTTTACTAGAAATTA--GAGCATTAAGCCCTGAA 1203

Qy 91 IleThrValTrpLeuAspLeuLeuLysProIleValLysGlnIleArgProLysHis 110
    |||||
Db 1204 CTGATCTCTTTTAAGAAAGAACAAATTCATTGGAACATGAATAAACGTTCCACCGAA 1263

Qy 111 ValVal-----ValLysPheValValLysPhePheProAspHisIleThrGlnLeuGln 128
    |||||
Db 1264 CTACTAGAAACTGTTTCTTTGACTTACGCTAACGCAAGAAAGAAATTAACCTTCACTAGA 1323

Qy 129 GluGluLeuThr-----ArgTyrLeuPheAlaLeuGlnValLysGln-----Asp 143
    |||||
Db 1324 CAGAAAGTAAATGCTGTGAAGCAAAATATTCATTCTGTG--GTGAACAAACGTTGGAT 1380

Qy 144 LeuAla---GlnGlyArgLeuThrCysAsnAspThrSerAlaAlaLeu---LeuIleSer 161
    |||||
Db 1381 TTGGCCGTGCAAGTAACTACTCTTTTAAATGATCATCTGGATACAGAAACGGCATCG 1440

Qy 162 HisIleValGlnSerGluIleGlyAspPheAspGluAlaLeuAspArgGluHisLeuAla 181
    |||||
Db 1441 CCGTTATCGCAAGTATGATATCATTTAAGAAATACTGCGAGCAAGCAAAATGCGTA 1500

Qy 182 LysAsnLysTyrIleProGlnGlnAspAlaLeuGlnLysArgLysIleValGluPheHis 201
    |||||
Db 1501 AATGAAAT-----GATCCCAAGCCATCTTTCGAAAGACTGCTGTAATTCAGTAAAT 1554

Qy 202 ---AsnHisIleGlyGlnThrProAlaGluSerAspPheGlnLeuGlu----- 217
    |||||
Db 1555 GTGAATGAGCTA-----CAAGAAAAAATGTGTAACCTTGAAATTCGATTCGG 1602

Qy 218 ---IleAlaArgAlaGluGluMetTyrGly----- 226
    |||||
Db 1603 ATTTTGCTGATTAATTAGAAATTTATGAGAAAAACAAGACAAGACCTTACAAAGGTA 1662

Qy 227 -----IleArgLeuHisProAlaLysAsp 234
    |||||
Db 1663 GAAACACAGACCTTAAAGAAAGCAAAAGATGCAATTAATGACTGGAACATCAATCGC 1722

Qy 235 ArgGluGlyThrLysIleAsnLeuAlaValAlaAsnThrGlyIleLeuValPheGlnGly 254
    |||||
Db 1723 AAGATGGAACCCGATTAATATATCCT-ATTAAAGAGCGGATTCCTTACAAATTCGTGGC 1781

Qy 255 Phe-ThrLysIleAsnAlaPheAsnTrpAlaValArgLysLeuSerPheLysArgGly 274
    |||||
Db 1782 ATCAACTGAGAAAAACAAGCCAAACAAATTCGTGCACTTATGAGAGACGCCCTGGA 1841

Qy 274 SarGpHeuLeuIleLysLeuArgProAsp-----AlaAsnSerAlaTyr 288
    |||||
Db 1842 AAAAGAAATTTAGAACTTGAGCGAAGAACTTTCATCTACAAATTCGAAATCTGCTAT 1901

Qy 288 rGlnAspThrLeu-----GluPheLeuMetAlaSerArgAspPheCysLysSerPheTr 306
    |||||

```

```

Db 1902 TATTCAAGATTTAAGAAAAAGCTGTGAATTTATTAAGAAATCAAAATGCGATAA----- 1956

Qy 306 rLysIleCysValGlnHisIleAlaPhePheArgLeuPheGluGluProLysProLysP 326
    |||||
Db 1957 -AAAACACTTTTAGAGAGATTTGAAAAATTTCAAGGGTTGGCAAGAGAAAAAGAA----- 2010

Qy 326 oLysProValLeuPheSerArgLysSerPheArgPheSerGlyArgThrGlnLysG 346
    |||||
Db 2011 -----CGTATGTTAGAGA 2024

Qy 346 nValLeuAspTyrValLysGluGlyGlnHisLysLys----- 358
    |||||
Db 2025 AGCAATAGACCCCTTAAAGCTGAACCTGCAACGAAACAAATCATGGCTTCATCATACAT 2084

Qy 359 ---ValGlnPheGlnArgLysHisIleSerLysIleHis-----SerIleArgSe 373
    |||||
Db 2085 ACATGTTGAAAGAAAGAAAGAGCCCTCAACTGACCTGACAGCTTACAGATCAAAATCAATC 2144

Qy 373 rLeuAlaSerGlnProThrGluLeuAsnSerGluValLeuGluGlnSerGlnIleSerTh 393
    |||||
Db 2145 TTTGAGATGAGAGATTTCCAAACTGAAAGAAAGAGACGCGACCTTATACCCAGAAAGA 2204

Qy 393 rSerLeuThrPheGlyGluGlyAlaGluSerProGlyGlyGlnSerCysArgArgGly 413
    |||||
Db 2205 ATCTTTAACT-----AGAGATTTGCAACAAATGCTGCAAGAAAAAAA 2246

Qy 413 sGluProLysValSerAlaGlyGluProGlySerHisProSerProAlaProArgArgSe 433
    |||||
Db 2247 AGAATTCAGATGAGCTAAAGAG-----TGTGAATTCAGCATTAATGAAATTAAT 2300

Qy 433 rProAlaGlnLysGlnAlaAspGlyAlaAlaSerAlaProThrGluGluGluGlu 453
    |||||
Db 2301 GGAATTTAGTTGCAAGAGAGGCGCAATACAAAGCCAAATC-----AAAGAAATGGA 2351

Qy 453 vValValLysAspArgThrGln-----GlnSerLysProGlnProProGlnProse 470
    |||||
Db 2352 AATCAGTTTGAAGAAATTAAGTTACGTTCCGACTTACATCAATCAAAAGATCAAA----- 2397

Qy 470 rThrGlySerLeuThrGlySerProHisLeuSerGluLeuSerValAsnSerGlnGly 490
    |||||
Db 2398 -----GAAATGAATCGATACGCTC 2417

Qy 490 yValAlaProAlaAsnValThrLeuSerProAsnLeuSerProAspThrLys---GlnAl 509
    |||||
Db 2418 ATGCAAGATTCGCAACTCAATGAGGCTCAAAATACCATGAGATGATACAGAAATGAAT 2477

Qy 509 aserProLeuIleSerProLeuLeuAsn---AspGlnAlaCysProArgThrAspArg 528
    |||||
Db 2478 GAACTCCTTAATACTGAATTAATCAATAGAAAGAAACAACAAATTTGAAATTCATCAGA 2537

Qy 528 uAspGluGlyArgArgLysArgPheProThrAspLysAlaTyrPhe-----Il 544
    |||||
Db 2538 AATAGAGAGCTTGATTAAGAACTAAGAAAAACCAATTCATATTAATTTCTTGATCA 2597

Qy 544 eAlaLysGluValSerThrThrGluArgThrTyrLeuLysAspLeuGluValIleThrSe 564
    |||||
Db 2598 AATATGATGATGCAAGACGTTAGAACCGACTTGAGAAAAAGATTAAGACAAATCCAAAGT 2657

Qy 564 rTrpPheGlnSerThrValSerLysGluAspAlaMetProGluAlaLeuLysSerLeuI 584
    |||||
Db 2658 TCAACTGAGAGATGCAAACTCCACAGATACAGGCTTTGAGAGATTAATTAATCTTCCG----- 2712

Qy 584 ePheProAsnPheGluProLeuHisLysPheHisIleThrAsnPheLeuLysGluIleGlu 604
    |||||
Db 2713 -----AATGAAATGCGCTGATGAAATTTGAAAAA 2741

Qy 604 nArgLeuAlaLeuTrpGluGlyArgSerAsnAlaGlnIleArgAspTyrGlnArgIleG 624
    |||||
Db 2742 TGAATAGCAAAAAACAAGAAATTAATGATGCTAAATTAATGAATGAGAGAAG----- 2796

Qy 624 yAspValMetLeuLysAsnIleGlnGlyMetLysIleLeuAlaHisIleLeuTrpLysHis 644
    |||||
Db 2797 -----GAAAAATGGCGGAG 2810

```

```

Qy      644 sserglualaleuglualaleugluanglylleysserargargluegluaenph 664
Db      2811 GGAAAGATGATCTCTCGTTTAAACGGAGAGCTGGTGAATTAAGCTTGGCAACCGAA 2870
Qy      664 ecyargphrphneglualeuglnlyvalcyetyrleuproleuanthrphleuaur 684
Db      2871 ACTAAAGAGAGTGATTCAT-----TTCTTCAACATACAGAAAAGTTAGAGAA 2921
Qy      684 gproleuuh1arglueuch1slyrlyeglnvalleuglnjrg----- 698
Db      2922 TGAGGTAGAGAGATCCAAAATAATGTAGAAAATAATGAAAATAAGAGCATTTGTCCA 2981
Qy      699 -leucyelyuh1enhrprofer----- 706
Db      2982 ACTGTGCAAAAAAAGAAATGCCCAATATGCATGCACAAATGAAGAAATAAGATCT 3041
Qy      707 -----H1ealaprphargphcyargalalaleualag1 719
Db      3042 ATCTGAATGGTATCCGCTTGAAGAGATCGGCTATTTGCCAA-----GCAGA 3092
Qy      719 u1lethrgluevalalaglnleuuh1eglythmerlleuymetgluaenpheln- 738
Db      3093 ATTAACAAAACCAAAATCCTCATTTGTATAGCAGCATGATTTACTGATTAACATGAAG 3152
Qy      739 -lyleuuh1egluelylyvhrleuuh1eglyleapnleuvalvalproglyar 758
Db      3153 GAAATGGATGAGGAAAAGCGGTATTAT-----GAAAG 3185
Qy      758 gglpheh1leargluegllyserleuyseryleuyserylyvsglyleuglnlaryme 778
Db      3186 AGAATTGAT-----TCTAATATTGA-----CAGACGAATC 3218
Qy      778 rphrphleuhrhenaervalleuyleuylrthserarglyleuthralaserang1 798
Db      3219 CTTCGCTGTGAAAATTCGATTTGATTAAGAGGTATGATGACACAGCAAAATATGCG 3278
Qy      798 nphelyvalh1eglylnleupro-----leuylrlymetthrllleglueglu 816
Db      3279 CGATTAAGATCATTTGAAGCTAGTTCTTTGTTTCAACTTACCTCATGAAGAAATTC 3338
Qy      816 uaprgl1trg1yvalproh1sylvleuthrleuargllylnarglinsertllele 836
Db      3339 TTTAGAGACGAAGTTAACCACTGC-----AAGAGCAATGGCATT 3380
Qy      836 lalalaserarglueglumetgllyetrvalgluaprlleglmetalaileas 856
Db      3381 CGTAAACAAAAGAACATAGTTTAGAAAATAAGCTGCAACGAACCCAGAC 3440
Qy      856 pleuhalaglylye 860
Db      3441 TTATCGGAGAG 3453

```

```

RESULT 15
US-11-217-529-173406
; Sequence 173406, Application US/11217529
; Publication No. US20060099612a1
; GENERAL INFORMATION:
; APPLICANT: SUNTORY LIMITED
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKAI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: S-38-285
; CURRENT APPLICATION NUMBER: US/11/217,529
; PRIOR FILING DATE: 2005-09-02
; PRIOR APPLICATION NUMBER: US 10/932,182
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: Patent version 3.3
; SEQ ID NO 173406

```

```

; LENGTH: 1617
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
; US-11-217-529-173406

Alignment Scores:
Pred. No.: 0.421 Length: 1617
Score: 136.00 Matches: 84
Percent Similarity: 37.0% Conservative: 69
Best Local Similarity: 20.3% Mismatches: 152
Query Match: 2.5% Indels: 108
DB: 7 Gaps: 18

US-09-555-342B-2 (1-1045) x US-11-217-529-173406 (1-1617)
Qy      473 serleuthrglyserpro--H1eulesergluleuervalaensergllyval 491
Db      28 ACTTGAAGGGGTGAGGCTCATATAGTAGTTCATTGTCCCGCAAAAGAACCTTGA 87
Qy      492 Alarproalaevalthr-----leuserproanleuserproasp--- 505
Db      88 GCGATTCTAAGCGTCAATGGTGTGTTACATTTACTGGCGGAATGCTATTTCCCTACGTGC 147
Qy      506 Thrlysglnlaserproleu1leserproleuenaerglalacyproarg--- 524
Db      148 ACAAGACGCT-----TTGTACTATTTCTGTCCAAAGCGT 189
Qy      524 ----- 524
Db      190 CTGAGCAACAGCAAGTTAATCTGAACCAATAGCGGCTGAACGTAATATATCA 249
Qy      525 -----Thrapargluarpgllyarg----- 532
Db      250 ATGGCATTGATGATGATGACGAAGAAAGCCATTTGACAACTCCGATCTAAAGAG 309
Qy      533 -----Arglyargphrprothrlyarplyvalrphlealaly 546
Db      310 GATTGGCAAACTTACCAAAAAGAAATTAAGACAGAGGCTATTAT----- 360
Qy      547 gluv1serthrthrgluargthrglyleuylusarpleugluval1lethsertrphe 566
Db      361 GAAGTTATATACACGGAAGAACCTTGTAAATCATTAAGAAATTAACAGACACATTC 420
Qy      567 glnerthrvalserlyvsgluaralamecproglualaleuys-----Ser 582
Db      421 ATGAAGACTTTGGCTGAACCAATATTTTACACAGATATTAAGAAAATTTATCAAG 480
Qy      583 leu1lephrproanphnegluvproleuuh1sylvrphen1erthrapnleuylg1ule 602
Db      481 CATGTTTTCACATATTAAGACATATATTCATGCAATAGACGTTTTCAGAGCTTTA 540
Qy      603 gluglnargleualaleutrgllyargseranlaaglnleargarvrgl1arg 622
Db      541 ACTGATAG-----CAAGAAATTCACACAGTTTGAGAGGA----- 576
Qy      623 l1leglyarvalmetleuylasnlleglnlymetuynhsyleualalanh1leuttr 642
Db      577 ATTGGGACATGTTTGAAGATTCATTCATTTTAAACCATTTGTTTCAATAGCGCT 636
Qy      643 luh1sserglualeuglualeugluynh1lylleysserargarglueglu 662
Db      637 TCAAGACCATATGCTAAGTACTTATTAACCCAAAGATGACATCCATTTTGGCT 696
Qy      663 Aanrhecyargarphneglualeuglnlyvalcyetyrleuproleuanthrphleu 682
Db      697 AGATTGATGACGATATGATAGCTCCTTACGTCAT---GGATATGATTCGTTTAA 753
Qy      683 leuargproleuuh1sarglueuch1slyrlyeglnvalleuglnjrglueglue 702
Db      754 TCTCAAGGTGTTCAAGACGATGATATGCTATTTGGCAAGAAATATGAAATCT 813
Qy      703 H1spropro-----Serh1ealaprphargphcyargalalaleu 717

```

```
Db      814 ACTGACCCCGAAGACAAAGTGAATTATGATTTAAGCAAGCCATGATGCATTA 873
Qy      718 AlagluilethrgumetValAlagluLeuHisgLyThrMetIlelysmetGluAsnPhe 737
      874 AGGGAATTTATGAAAAGATTTGATCAAGCAAGTGTCTGCA----- 915
Db      738 GlnIysLeuHisgIleuLys-----LysAspLeuIleGlyIleAspAsnLeuValVal 755
      916 CAAAGATAGACATGATGTAAATTTGTGAACAAAGATTCTTTCAAAAC----- 966
Qy      756 ProGlyArgGluPheIleArgLeuGly----- 764
      967 -----GAAATACGTTAAATTTAGATTAATGACGAAAGAGAAAAATTAAACATGAA 1017
Qy      765 -----SerLeuSerIysLeuSerGlyLysGlyLeuGlnGlnArgMet 778
      1018 GGTATTAATCAACGAAAGAGTTGTCAAAATCGAGCGAAACAGTTGTGTGACATACAA 1077
Qy      779 PhePheLeuPheAsnAspValLeuLeuTyr---ThrSerArgGlyLeuThrAlaSerAsn 797
      1078 TTTTATCTTCTAGATATATATGCTCTTATTTTAAAGCCAAAGCAGTTAATAATGCGAC 1137
Db      798 GlnPheLysValHisgIyGln-----LeuProLeu 807
      1138 CAGCATAGGATTCCAAGGCCCATTCATTCGCATTG 1176
```

Search completed: May 20, 2006, 04:15:02
Job time : 156 secs


```

Qy 20 SerGlyIleSerThrLeuGluArgGlyGlnIys----- 30
Db 214 GCGCTCCGCCACCTGTGGCCGAGAGGACAGACATGAAATGATCCATTTGGAGCT 273
Qy 31 ProProThrProSerGlyLysLeuValSerIleValIleGlnMetLeuAspArgThr 50
Db 274 CCACGCTCAAAATCCCTGACGACGAG--ATGAGATGACACATCCGCTGTGTGAAGACTCG 332
Qy 51 GlnGlnAlaPheGlnValProGlnArgAlaProGlyLysValLeuLeuAspAlaValCys 70
Db 333 GAGATCTTCGCCACATCCAGACGAGAAACCAAGGGGAGCTTTTCATTTGACACATCTGCG 392
Qy 71 AsnIleLeuMetLeuValGlnGlyAspArgPheGlyLeuGlnPheProAsnIleLysLys 90
Db 393 AACCTACACAGCTGCTGGAGAAAGACTATTGGCATGCTGCTGTGGACCCGAGAAAG 452
Qy 91 IleThrValIlePheAspLeuLeuLysProIleValLysGlnIleArgArgProLysIle 110
Db 453 CAAGGCACTGGCTTGAACCTAACAGATCCATCTTCAAGCAAAATGAAACTCATCCASCA 512
Qy 111 ValValValLysPheValValLysPhePheProPheAsnIleSerGlnLeuGlnGlu 130
Db 513 TACACCAATGCTTAAAGTAAATTTCTACCATATGAACTTGAAGATTAAAGAGAG 572
Qy 131 LeuThrArgTyrLeuPheAlaLeuGlnValLysGlnAspLeuAlaGlnGlyArgLeuThr 150
Db 573 CTCACCAAGATACCTTTATACCTTCAAGATTAAGGACATTTTTCATGGCCACTGCTG 632
Qy 151 CysAsnAspThrSerAlaAlaLeuLeuIleSerHisIleValGlnSerGlnIleGlyAsp 170
Db 633 TGCTCCCTTTCTGATGCTGCTTACCTGAGTCCGTGATTTGTTCAAGCTGAGCTTGAT 692
Qy 171 PheAspGlnAlaLeuAspArgGlnHis-----LeuAlaLysAsnLysTyrIle 186
Db 693 TACGAT-----CCTAATGACATCTCGAATTAACATGATGAGTTGGAATTTTC 743
Qy 187 ProGlnGln--AspAlaLeuGlnAspLysIleValGlnPheHisAsnHisIle--- 204
Db 744 CCACAGCATGACAGAGCTGGAAGAAATATGGAATCTCAATAAATGAACTCAGG 803
Qy 205 GlyGlnThrProAlaGlnSerAspPheGlnLeuLeuGlnIleAlaArgArgLeuGlnMet 224
Db 804 GGGGAGAGCCACAGCTGTGATTTAACTTGTCTCTGAAAGCTCACACTTTGGAAGACC 863
Qy 225 TyrGlyIleArgLeuHisProAlaLysAspArgGlnGlyThrLysIleAsnLeuAlaVal 244
Db 864 TACGGGGGATCTCTCACCCATGCAAGGATTCACACGACACACACATTTTATGAGATTC 923
Qy 245 AlaAsnThrGlyIleLeuValPheGlnGlyPheThrLysIleAsnAlaPheAsnTyrAla 264
Db 924 ACAGCTGCAAGCTTTGTGTCTTTCAGGGAAATGAAGAAATCCATTTGATTAATGCGCA 983
Qy 265 LysValArgLysLeuSerPheLysArgLysPheLeuIleLysLysLysArgProAspAla 284
Db 984 GATGTCTGCAAAATGGAATTGAAGGAGACATTTATGATGTGACACCCAGAGAGAG 1043
Qy 285 AsnSerAlaTyrGlnAspThrLeuGlnPheLeuMetLaseArgAspPheCysLysSer 304
Db 1044 AAAAAAGCC-----ATGTTGGCATTTCAATCTTCAACACCAAGCTGCTGAAACAT 1094
Qy 305 PheTyrLysIleCysValGlnHisAlaPhePheArgLeuPheGlnGluProLysPro 324
Db 1095 CTTTGGAGAGGTGAGTGAAGAAACAGAGCTTTTATATAGTATGCAAAATCCAGTCAGATC 1154
Qy 325 Lys-----ProLysProValLeuPheSerArgLysSerPheArgPheSerGlyArg 342
Db 1155 AAGCTGTATACACACAGACAAATATTTTAAAGAAATGATATTTTCATATAGTGGAGAA 1214
Qy 343 ThrGlnLysGlnValLeuAspTyrValLysGlnGlyGlnHisLysValGlnPheGln 362
Db 1215 GTTGGCAAGAGGTGTGAG-----GCCAGTTCCAAAGATCCAG----- 1253
Qy 363 ArgLysHisSerLysIleHisSerIleArgSerLeuAlaSerGlnProThrGln----- 380

```

```

Db 1254 AGGAGCTCTCTAGGTGACAGAGCAACATTTACTCAGAGCCGACGTTCCACTCTTG 1313
Qy 381 -----LeuAsnSerGlnValLeuGln-----GlnSerGln 390
Db 1314 AACAAACAGCTATCATTAACATGAACCCCTGACGCTGCTTCTTCCCGACGAGG 1373
Qy 391 GlnSerThrSerLeuThrPheGlyGlnGlyAlaGlnLysProGlyGlnSerCysArg 410
Db 1374 CAGAGAAAGAACTTCTCTGGGTGAGGGGTTCATGGCT----- 1415
Qy 411 ArgGlyLysGlnProLysValSerAlaGlnLysProGlySerHisProSerProAlaPro 430
Db 1416 -----AAAGAGAGAAACATTTCTGCTCCCTTGATGTCGACGTCGCCAGAGGAGCGCC 1469
Qy 431 ArgArgSerProAlaGlnLysLysGlnAlaAspGlyAlaAlaSerAlaProThrGlnGlu 450
Db 1470 CCGGAG-----TATGAGATGCCCTGATGAA 1496
Qy 451 GlnGlnGlnValValLysAspArgThrGlnGlnSerLysProGlnProProGlnProSer 470
Db 1497 GAGAGAGATTAATAAAGAA-----GAACTTTA 1526
Qy 471 ThrGlySerLeuThrGlySerProHisLeuSerGlnLeuSerValAsnSerGlnGly 490
Db 1527 ACC-----ATCTGGAATGATGTAACAAACCAAGTGCAGC 1562
Qy 491 ValAlaProAlaAsnValThrLeuSerProAsnLeuSerProAspThrLysGlnAlaSer 510
Db 1563 CTGCTCCCAACCCCTGTG-----GATGACGATGAGATTGAC 1598
Qy 511 ProLeuIleSerProLeuLeuAsnArgGlnAlaCysPro-----ArgThrAspAsp 527
Db 1599 ATGCTCTTGAC-----TGCTCTTCTAGCGCTTGAGTTGGAAGA 1637
Qy 528 GlnAspGlnGlyArgArgLysArgPheProThrAspLys--AlaTyrPheIleAlaLys 546
Db 1638 GAAGACACAGATTCATTTGAGATGATGGAAGCAGATGAAGAAACGCTTTTATGATGCTGAA 1697
Qy 547 GlnValSerThrThrGlnArgThrTyrLeuLysAspLeuGlnValIleThrSerThrPhe 566
Db 1698 GAAGAGAG-----CTGAAGAGAGCTGCGCTGCTTGTGTGTGAGC 1739
Qy 567 GlnSerThrValSerLysGlnAspAlaMetProGlnAlaLeuLysSer----- 582
Db 1740 TACGACATTTCTGACTGGCCATTTGCGGTGAACCCACTGCTCAAGATTTTCCAGGCTC 1799
Qy 583 -----LeuIlePheProAsnPheGlnProLeuHisLys 593
Db 1800 CTTGTGGTGGGCTGGAGACTGCTGCTTTGTGATTTCCCTGCTCTCTCTCTTTTGAG 1859
Qy 594 -----PheHisThrAsnPheLeuLysGlnIleGlnGln 604
Db 1860 TCAGGTATGATCTCTCTTTTATGCGAAATCGCGAG 1898

RESULT 2
US-11-412-475-7
; Sequence 7, Application US/11412475
; GENERAL INFORMATION:
; APPLICANT: The Salk Institute for Biological Studies
; APPLICANT: Novartis, Iliris
; APPLICANT: Verma, Indar M.
; APPLICANT: Lockhart, David J.
; APPLICANT: Barlow, Carolee
; TITLE OF INVENTION: ANXIETY RELATED FACTORS AND METHODS OF USE
; FILE REFERENCE: 7518-70020-02
; CURRENT APPLICATION NUMBER: US/11/412,475
; PRIOR APPLICATION NUMBER: 2006-04-26
; PRIOR FILING DATE: 2005-04-27
; NUMBER OF SEQ ID NOS: 115
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 7

```

LENGTH: 2332
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (275) .. (2026)
US-11-412-475-7

Alignment Scores:
Pred. No.: 1.4e-24 Length: 2332
Score: 535.00 Matches: 197
Percent Similarity: 40.7% Conservative: 133
Best Local Similarity: 24.3% Mismatches: 288
Query Match: 9.8% Indels: 193
DB: Gaps: 30

US-09-555-342b-2 (1-1045) x US-11-412-475-7 (1-2332)

Qy 7 ArgProThrProGlySer-----ArgLeuGluAlaProGluAAsnSerGlyIleSer 23
Db 32 CGCCCAAGTCCCGCGCGCTTCCCGCGCGCTCCCGCGCGTGAAGTCTCCCGCGCTTTC 91
Qy 24 ThrLeuGluArg-----GlyGlnLysProProProThrPro 35
Db 92 CGCCTTCTCGCTCCCGCGCGCGCGGCTTCCGCGCGCGTCCCGGACCCGCGCGG 151
Qy 36 Ser----- 36
Db 152 AGCTGCTGCTCAACCTCGCAACCGCGCGCAATAAATTATGATCGGCTTGAGTG 211
Qy 37 -----GlyLysLeuValSer----- 41
Db 212 GAGGCGCGCGCTCGCGCGCTCGGTGAGTCTTCTCTCGCTGCGAGACACACAAAC 271
Qy 42 ----- 42
Db 272 GACATGGGCTGTTTCTCGCGCTTCCGGAAGATTCTACGAAAGTTTCTCTGAT 331
Qy 49 AspThrGlnGluAlaPheGluValProGlnArg-----AlaProGlyLys 63
Db 332 GAATCCCAAGTTAACCTTACACCCACGACGAGGCGCATCAAGAGTCAACGAAAGTTCC 391
Qy 64 ValLeuLeuAspAlaValCysAsnHisLeuAsnLeuValGluLysAspThrPheGlyLeu 83
Db 392 GTTCTCTTGACACGATTCCTCATCACTTAACCTTGGAGATATGATTTTGGGCTA 451
Qy 84 GluPheProAspHisValLysIleThrValTrpLeuAspLeuLeuLysProIleValLys 103
Db 452 CGTTACTGTACAGAAAGCATCAGACGATTTGGCTGATCTGGCAAAACCTTGCTGAA 511
Qy 104 Gln-----LeuArgArgProLysHisValValLysPheValValLysPhePhe 120
Db 512 CACAAAGATGATCAACATGACATGACCTTCATATCTTGTATTTGGTATTAATTTCTAT 571
Qy 121 ProProAspHisThrGlnLeuGlnGluGluLeuThrArgTrpLeuPheAlaLeuGlnVal 140
Db 572 GCTGAAGTCCATGATTAACATAAGAAATTAACACATATCACTTTTCTTGACGTG 631
Qy 141 LysGlnAspLeuAlaGlnGlyArgLeuThrCysAsnAspThrSerAlaLeuLeuIle 160
Db 632 AAGCAAGATGCTCTTACGCGCGCTGCTGCCCTCCCAACATGCGTCAAGTGA 691
Qy 161 SerHisValLeuAlaGlnSerGluIleGlyAspPheAspGluAlaLeuAspArgLys 179
Db 692 GCGATGACATCCAGTCCGAGCTGGAGATTAAGCCATATTAACATATGACAGATAT 751
Qy 180 LeuAlaLysAsnLysValTrpPro-----GlnGlnAspAlaLeuGluAspIleValGln 198
Db 752 GTATCTGAGTACCGGTTTCTGATCATGAAGAAAGAACTTGAAGAAAGCCATGAAG 811
Qy 199 PheHisValAsnHisValGlyGlnThrProAlaGlnSerAspPheGlnLeuLeuIle 218
Db 812 ATTCAATAAACTGTAATGAGTCAAGATTCCTTCTGAGCTGAGCTGAATTACTTGAGACT 871

Qy 219 AlaArgArgLeuGluMetTrpGlyIleArgLeuHisProAlaLysAspArgGluGlyThr 238
Db 872 GCCAAATCCCTGGAGATGATGGCTTGACCTTCATCCGCTTATGGAGAAAACAAGTCT 931
Qy 239 LysIleAsnLeuAlaValAsnThrGlyIleLeuValPheGlnGlyPheThrLysIle 258
Db 932 GAGTATTTCTTAGATTAACCTCCGTTGGTGTGTGTGTCAGAAATTAAGAAAGTGG 991
Qy 259 AsnAlaPheAsnTrpAlaLysValArgLysLeuSerPheLysValGlyAspPheLeuIle 278
Db 992 GGGAAAGTATTTCTGCTCCGATTTACAAAGGTTCACTTCAAGAGACTCAATTGAACTC 1051
Qy 279 Lys-----LeuArgProAspAlaAsnSerAlaTrpGlnAspThrLeuGluPheLeuMetAla 297
Db 1052 AGAGTACTGGGAAAAGATTTGATC-----GAAACCTCATTTCTTTTGAAGCTCGG 1102
Qy 298 SerArgAspPheCysLysSerPheTrpLysIleCysValGluHisAlaPhePheArg 317
Db 1103 AGTAAACTGCTTGCAGACACCTTGAGAGTGCAGTGGAAACATCATATCTTTTATGA 1162
Qy 318 LeuPheGluGluProLysProLysProLysProValLeuPheSerArgGlySer----- 335
Db 1163 ATGCCAGAAAATGATCAATTCATCTGCAAGAAACTCAGACAGTTTGATCCATAGCT 1222
Qy 336 ---SerPheArgPheSerGlyArgThrGlnLysGlnVal---LeuAspTrpValLysGlu 353
Db 1223 TATTAAGCACCGCTACAGTGGAGGACACGCTTGGCAATATGACCGAGATCTTTCTATTCAG 1282
Qy 354 GlyLysHisLysLysValGlnPheGluArgLysHisSerLysIleHisSerIleArgSer 373
Db 1283 CTTCGCCGCGCTGATCAAGATGTGACAAAGTGGAAAGCAAGATTACCTTAACGATA 1342
Qy 374 LeuAlaSerGlnProThrGluLeuAsnSerGluValLeuGlnGlnSerGlnSerThr 393
Db 1343 GCACAAACACAGCCGCTGATCAAAACACC-----ATCAGTAGGATTAACGTCA 1390
Qy 394 SerLeuThrPheGlyGluGlyAlaGlnSerProGlyGlyGlnSerCysArgArgGlyLys 413
Db 1391 AACATG-----GAAATGCAAGAAATGAAGA----- 1417
Qy 414 GluProLysValSerAlaGluGluProGlySerHisProSerProAlaProArgArgSer 433
Db 1418 ACATTAATAATTATGCA-----CCTTACACGAGAAAAGCTTTAAG 1459
Qy 434 ProAlaGlnLysGlnAlaAspGlyAlaAlaSerAlaProThrGluGluGluGlu 453
Db 1460 AAAGCAAGATGAATAATAGCCCTGAT----- 1486
Qy 454 ValValLysAspArgThrGlnGlnSerLysProGlnProGlnProSerThrGlySer 473
Db 1487 -----ACCCAAAGCAAAATCTGATCAGCCGTGGAGAAAGAAATGGCCCC 1531
Qy 474 LeuThrGlySerProHisLeuSerGluLeuSerValAsnSerGlnGlyValAlaPro 493
Db 1532 CAGAGTGAAGCTTCAAT-----TCTCCC 1555
Qy 494 AlaAsnValThrLeuSerProAsnLeuSerProAspThrLysGlnAlaSerPro----- 511
Db 1556 AGTATGCACTAATGCGCCAAAGTTC---CCTTAACAGCGCTGCCCAAAACCTCTCTGT 1612
Qy 512 -----LeuIleSerProLeuLysAsnArgGlnAlaCysProArgThrAsp 526
Db 1613 GGAAGTGAATGATTTCTGTACACCTGTGTGAGAGAGAAAGCCATTAACAGTGGTGA 1672
Qy 527 AspGluAspGluGlyArgArgLysArgPheProThrAspLysAlaTrpPheIleAlaLys 546
Db 1673 GATTCAGATCTTAAAGAAAGAGAGGATCACTTACAGCTGTAAACACAGAGTGTAGT 1732
Qy 547 GluValSerThrGlnGluGlyThrTrpLeuLysAspLeuGluValIleThrSerTrpPhe 566
Db 1733 GAATCAGAAATTTCTAATAGAAACACCGGAAAAG----- 1768

```

Qy 567 GlnSerThrValSerTysGluAspAlaMetProGlnAlaLeuTyrSerLeuIlePhePro 586
      ::::: ::::: ::::: ::::: :::::
Db 1769 AGAAGACAGATACGGACAGAGATGATGATGTTGATTCAGCG-----CCT 1813
Qy 587 AsnPheGluProLeuHisIlePhePheHisThrAsnPheLeuTyrGluIleGlnGlnArgLeu 606
      ::::: ::::: ::::: ::::: :::::
Db 1814 CAGTGGGAAGCTGTA-----TTTAGAGAGACAAAGAAAAAAMC 1852
Qy 607 AlaLeuTrpGluGlnArgSerAsnAlaGlnIleLeuArgAspTyrGlnArgIleGlyAspVal 626
      ::::: ::::: ::::: ::::: :::::
Db 1853 CAACCCACCCCAACAGACGGCGATCCAGACACAGATCTCTTGAGAGAGCCCGGATATC 1912
Qy 627 MetLeuTyrAsnIleGlnGlyMetLeuHisLeuAlaHisLeuTrpLysHisSerGlu 646
      ::::: ::::: ::::: ::::: :::::
Db 1913 CAACAAAGAAAGAA-----GAGTTATGGAAACACATTCAA 1945
Qy 647 AlaLeuGlnAlaLeuGlnValAsnGlyIleLeuSerSerArgArgLeuGlnAsnPheCysArg 666
      ::::: ::::: ::::: :::::
Db 1946 AAGAACTTGATGATCCATCCGAGATTGTCGAGAA----- 1981
Qy 667 AspPheGlnLeuGlnIleValIleCysTyrLeuProLeuAsnThrPheLeuLeuArgProLeu 686
      ::::: ::::: ::::: ::::: :::::
Db 1982 -----CAATTAAAGAGATTCATACCTAAATAGAGAC-----ACAAGGTGACCC 2029
Qy 687 His-ArgLeuMetHisTyrTyrGlnValLeuGlnArgLeuCysLysHisHisProPhe 706
      ::::: ::::: ::::: ::::: :::::
Db 2030 AATCCGCATCAGCAT-----TCTCATTCGCCACGAG 2062
Qy 706 HisAlaAspPheArg-----AspCys-----ArgAlaAlaLeuAlaGlnIle 720
      ::::: ::::: ::::: ::::: :::::
Db 2063 TTACCGCGCATGATGCGAGGTCCAGTGTTCAGATGGGAGGAGATCAGTTCTCTCGAAGT 2122
Qy 720 e-----ThrGlnMetValAlaGlnLeu 727
      ::::: ::::: ::::: :::::
Db 2123 GAATTCAAAACAGATCTTGTAACCACTT 2153

```

```

RESULT 3
US-11-412-475-5
/ Sequence 5, Application US/11412475
/ GENERAL INFORMATION:
/ APPLICANT: The Salk Institute for Biological Studies
/ APPLICANT: Hovavata, Iliris
/ APPLICANT: Verma, Indar M.
/ APPLICANT: Lockhart, David J.
/ APPLICANT: Barlow, Carolee
/ TITLE OF INVENTION: ANXIETY RELATED FACTORS AND METHODS OF USE
/ FILE REFERENCE: 7518-70020-02
/ CURRENT APPLICATION NUMBER: US/11/412,475
/ CURRENT FILING DATE: 2006-04-26
/ PRIOR APPLICATION NUMBER: US 60/675,604
/ PRIOR FILING DATE: 2005-04-27
/ NUMBER OF SEQ ID NOS: 115
/ SOFTWARE: PatentIn version 3.3
/ SEQ ID NO 5
/ LENGTH: 2487
/ TYPE: DNA
/ ORGANISM: Mus musculus
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (52)..(1716)
/ US-11-412-475-5

```

```

Alignment Scores:
Pred. No.: 5,64e-22 Length: 2487
Score: 492.50 Matches: 163
Percent Similarity: 45.5% Conservative: 99
Best Local Similarity: 28.3% Mismatches: 216
Query Match: 9.0% Indels: 98
DB: 7 Gaps: 21

```

```

US-09-555-342B-2 (1-1045) x US-11-412-475-5 (1-2487)

```

```

Qy 43 LysIleGlnMetLeuAspAspThrGlnGlnAlaPheGlnValProGlnArg----- 59

```

```

Db 91 GAAGTTTGTCTCTGGACGAGTCCAAAGCTCAAGCTTCACCAAGCAGCAGCGGCGATCAAG 150
      ::::: ::::: ::::: ::::: :::::
Qy 60 -----AlaProGlyLysValLeuLeuAspAlaValCysAsnHisIleAsnLeuValGlu 77
      ::::: ::::: ::::: ::::: :::::
Db 151 AAGTCAACAAAGAGCTGTGTTCCTTGACAGCGTTTCCGTCACTAAACCTCTGTGGAG 210
Qy 78 GlyAspTyrPheGlyLeuGlnPheProAspHisIleLysValIleThrValIleTrpLeuAspLeu 97
      ::::: ::::: ::::: ::::: :::::
Db 211 ATGATATTATTTGGGCGGTGAGTCTGTGACAGAAAGCATCAGACATTTGGTGGATCCC 270
Qy 98 LeuLysProIleValLysGln-----IleArgArgProLysHisValValLys 114
      ::::: ::::: ::::: :::::
Db 271 GCAAAACCTTCAGAAACAAAGAGCTGATCAACAGCTGACCTCATATCTTGTAT 330
Qy 115 PheValValLysPhePheProProAspHisThrGlnLeuGlnGlnGlnLeuThrArgTyr 134
      ::::: ::::: ::::: ::::: :::::
Db 331 TTGGATTAAATTTCTTGCTGAAGATCCATGTAACTCAAGAAAGAAATACCAAGATAT 390
Qy 135 LeuPheAlaLeuGlnValLysGlnAspLeuAlaGlnGlyArgLeuThrCysAsnAspThr 154
      ::::: ::::: ::::: ::::: :::::
Db 391 CAGTTTTCCTTACAGGTGAAGCAAGATGCCCTCAGGGCGGCGCTTCCTCCGTCGAT 450
Qy 155 SerAlaAlaLeuLeuIleSerHisIleValGlnSerGluIleGlyAspPheAspGluAla 174
      ::::: ::::: ::::: ::::: :::::
Db 451 ATTGCTGCTCAGATGGGGGCGCATATGCCATCCAGCGCGAGACGACGACCCCATAC 510
Qy 175 LeuAspArgGlu---HisLeuAlaLysAsnLysTyrIlePro---GlnIleAspAlaLeu 192
      ::::: ::::: ::::: ::::: :::::
Db 511 AAGCACCTCGGGGAGTGTGTGTGAGATACCGGTGTGTCTGTATCGAAGAAAGAGCTG 570
Qy 193 GluAspLysIleValGlnPheHisIleAsnHisIleGlyGlnThrProAlaGlnSerAsp 212
      ::::: ::::: ::::: ::::: :::::
Db 571 GAAGAGCCATGAAGAGATTCATAAACCTTAAGGTGTCAGGCTCTTCGAAAGCTGAG 630
Qy 213 PheGlnLeuLeuGlnIleAlaArgArgLeuGlnMetTyrGlyIleArgLeuHisProAla 232
      ::::: ::::: ::::: ::::: :::::
Db 631 CTGAATTAATCTTGAGAGCTGCCAAATCCGGAAGATGATGTGTGACCTCCATCCCTGTC 690
Qy 233 LysAspArgGlnGlyThrLysIleAsnLeuAlaValAlaAsnThrGlyIleLeuValPhe 252
      ::::: ::::: ::::: ::::: :::::
Db 691 TATGAGAAATAATAGTCCGAGTACTTCTTGAGCTTCATCGGAGTGTGCTAT 750
Qy 253 GlnGlyPheThrLysIleAsnAlaPheAsnTrpAlaLysValArgLysLeuSerPheLys 272
      ::::: ::::: ::::: ::::: :::::
Db 751 AAGATATAAAGAGAGGGAGAGTATTTCTGCCCTCGCATCAAGAGGACCTTCAAG 810
Qy 273 ArgLysArgPheLeuIleLys---LeuArgProAspAlaAsnSerAlaTyrGlnAspThr 291
      ::::: ::::: ::::: :::::
Db 811 GAAACCCAGTTTGAGCTCAGAGCTTGCGGAAAGAGCTGTAAAT-----GAAACCTCA 861
Qy 292 LeuGlnPheLeuMetAlaSerArgAspPheCysLeuSerPheThrLysIleCysValGlu 311
      ::::: ::::: ::::: ::::: :::::
Db 862 TTCTTTTTCCTTAACTCGAAGCAAAAGCTCTTGCAAGCACTCTGAAATACAGCGTGAAG 921
Qy 312 HisHisAlaPhePheArgLeuPheGlnGlnProLysProLysArgProLysProValLeuPhe 331
      ::::: ::::: ::::: :::::
Db 922 CACCATACGTTTTCAGAAATGCCGACACAGATCAAAATTCATTATCAAGAAACCTAGC 981
Qy 332 SerArgGlySer---SerPheArgPheSerGlyArgThrGlnLysGlnVal---LeuAsp 349
      ::::: ::::: ::::: :::::
Db 982 AAGTTTGGGTCCATAAGTTATAACATCGGTACAGACAGCTTTCGAAAGAGCGGAGAT 1041
Qy 350 TyrValLysGlnGlyGlnHisLysLysValGlnPheGlnLysArgLysHisSerLysIleHis 369
      ::::: ::::: ::::: :::::
Db 1042 CTTTCTATTCACCTTCCCGGCCCAATCAAGACGTGTGTGAAGTGTCAAGACAGACTTAC 1101
Qy 370 SerIleArgSerLeuAlaSerGlnProThrGlnLeuAsnSerGlnValLeuGlnGlnSer 389
      ::::: ::::: ::::: :::::
Db 1102 CCGAAGAGGTATACAGACAGCTCAGCCCATCGATCAAAAGAAC-----ATCAATCGGATA 1155
Qy 390 GlnGlnSerThrSerLeuThrPheGlyGlyLys-----AlaGlnSerPro 404
      ::::: ::::: :::::

```

```

Db      1156 ACTGAAACAGGAAACCGGAGATGAGAAACCAAGATCATTCGCTCCCA 1215
Qy      405 GlyGlnSerCybArgArgGlyValGlnProValSerAlaGlyGlnProGlySer 424
Db      1216 ---GTGAAAGCTTTTAAAGAAAGCAAG-----AAT 1242
Qy      425 HisProSerProAlaProArgArgSerProAlaGlyAlaenlyGlnAlaSerGlyAlaAla 444
Db      1243 GAAAGACGCTTGACCCCAAGAAC-----AAATCT 1275
Qy      445 SerAlaProThrGlnGlnGln-GlnGlnValValSerAspArgThr----- 459
Db      1276 CATGACACCTTGGAAGAAACGAGCCGAGTGACTATACACTCTCCAGTACCGC 1335
Qy      460 -----GlnGlnSerlyProGlnPro---ProGlnProSerThrGlySerLeuThrGln 476
Db      1336 ACTAAGTACCAAGATTCTCCCTGTGCGCGGACGCAACTCTCTGCGGACGACACAT 1395
Qy      476 ySerProHis-----LeuSerGlnLeuSerValAs 486
Db      1396 GACTCTTCACAACTGATGAGCGGAAGAAAGCTCAACAAGTGGGAGAGACTCAGATCTC 1455
Qy      486 nSerGlnGlyGlyValAlaProAlaAsnValThrLeuSerProAlaenLeuSerProAspThr 506
Db      1456 AAGCAGAGAGAGAGGTCAAGCTCAGCTCAGCTGTAAACA----- 1489
Qy      506 rLyGlnAlaSerProLeuIleSerProLeuLeuAsnAspGlnAlaCybProArgThrAs 526
Db      1490 ---CAAGCACTGTGATGAGTCAAGAAATCTTAAACA----- 1522
Qy      526 pAspGlnAspGlnGlyArgArgGlyPheProThrAspIleValTyrPheIleAlaIle 546
Db      1523 ---GAGAACACCGGAAAGAAAGAAAGAAACCCCGCAGAGAAATGATAGTTCACCTCG 1578
Qy      546 gLysValSer-----ThrThrGlnArgThrTyrLe 556
Db      1579 GGGGCTCAGTGGGAACCGGTGTTAAGAGACAGAAAGAAAGCAAGCGGACCGCTTAAC 1638
Qy      556 u-----LysAspLeu-----GlnValIleThrSer 564
Db      1639 AACCGAGATCCAGACACAGATCTCGCTCAAGAAAGTCTGACATCC 1684

```

RESULT 4

```

US-10-448-162A-1
; Sequence 1, Application US/10448162A
; GENERAL INFORMATION:
; APPLICANT: Hart, Matthew J.
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES RELATED TO A GUANINE
; FILE REFERENCE: ONYX 1023.DIV2
; CURRENT APPLICATION NUMBER: US/10/448.162A
; CURRENT FILING DATE: 2003-05-27
; PRIOR APPLICATION NUMBER: 09/865,960
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 08/943,768
; PRIOR FILING DATE: 1997-10-06
; PRIOR APPLICATION NUMBER: 60/029,979
; PRIOR FILING DATE: 1996-11-06
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1
; LENGTH: 3150
; TYPE: DNA
; ORGANISM: Human p115 GEF-Rho
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (55) .. (2790)
; US-10-448-162A-1

```

Alignment Scores: 0.00555 Length: 3150
 Score: 180.50 Matches: 187
 Percent Similarity: 34.6% Conservative: 125

```

Best Local Similarity: 20.8% Mismatches: 351
Query Match: 3.3% Indels: 239
DB: 6 Gaps: 35
US-09-555-342B-2 (1-1045) x US-10-448-162A-1 (1-3150)
Qy      142 GlnAspLeuAlaGlnGlyArgLeuThrCybAsnAspThrSerAlaIleLeuLeuIleSer 161
Db      58 GAAGCTTCGCGCCAGAGGGGCGGCTCCCAAGCCCTCCGCGCTGGCTGTCCCTC 117
Qy      162 HisIleValGlnSerGlnIleGlyAspPheAspGlnIleAlaSerAspArgGlnIleAla 181
Db      118 AGCATTCATGGGCGTGAAGATGAGATTTTGAACAAGACCTGAGACAACTCAGAAAG 177
Qy      182 LysAsnlysTyrIleProGlnGlnAspAlaLeuGlnAspIleValGlnPheHis 201
Db      178 CAAAGAC----- 183
Qy      202 AsnHisIleGlyGlnThrProAlaGlnSerAspPheGlnLeuGlnIleAlaArgArg 221
Db      184 -----AGCCAGTTCCAGAGCTCGAGCAGTGAAGCG 216
Qy      222 -----LeuGlnMetTyrGlyIleArgLeuHisProAlaIleAspArgGlnGlyThrIlys 239
Db      217 CGCCAGCCGACCATCATGAGCCCTCGTGCAGACAGTGGCC----- 255
Qy      240 IleAsnLeuAlaValAlaAsnThrGlyIleLeuValPheGlnGlyPheThrIleAsn 259
Db      255 ----- 255
Qy      260 AlaPheAsnTrpAlaIleValArgIleLeuSerPheLysArgIlePheLeuIleLys 279
Db      256 -----CTGCAAGTTTGAAGCCAGAGACCCCTGCTTGCCTG 288
Qy      280 LeuArgProAspAlaAsnSerAlaTyrGlnAspThrLeuGlnIlePheLeuMetAlaSerArg 299
Db      289 CTGCATGCCGACATGCTGGCTCACTG---GGCCCAAGAGGACCAAGAGCCCTTCTCG 345
Qy      300 AspPheCybLysSerPheThrIleIleCybValGlnHisAlaPhePheArgLeuPhe 319
Db      346 GACTTTCACCAACACTTC-----CTGAGAGAGACAGGCGTTCCTCGGCTG--- 390
Qy      320 GlnGlnProLysProLysProLysProValLeuPhe-----SerArgLysSerPhe 337
Db      391 -----CCGGTCCCTCCCAAGCTGCTTTGAATTCAGCCGACATGAGGCTGAC 438
Qy      338 ArgPheSerGlyArgThrGlnIleValLeuAspTyrValIleGlnGlyIleHisLys 357
Db      439 CTGATCTCCGAGATGTCAGCGCGGTGCGTGCAGAGAGGTGGTGAAGC-----CAG 492
Qy      358 LysValGlnPheGlnArgIleHisSerIleHisSerIleArgSerLeuAlaSerGln 377
Db      493 CAGGTACCCGTGGCGCGAGCTGAGAGACTTCGTTCCAGACGCGCTCATGAGGATGACG 552
Qy      378 ProThrGlnLeuAsnSerGlnValLeuGln-----GlnSerGln 390
Db      553 CCTGGAGACAGAGAGCTGCCCACTGAGAGCTGGTGGGCGGAGCCAGACGCTAC 612
Qy      391 GlnSerThr-SerLeuThrPhe-GlyGlnIleGlnIleGlnSerProGlyGly----- 406
Db      613 GAGGCGCGGAGCGGACGTCGTGCGAGCGGTCTCATGACCTGAGAGAGATCAACAT 672
Qy      407 -----GlnSerCybArgArgGlyLysGln----- 414
Db      673 ACCATCTTACCGACGAAGAAAGATGCTGCGTGTCAACGCGCATTTGGCTGTACATG 732
Qy      415 -----ProL 416
Db      733 CGCCACTTGGGCGTGGCGGACCAAGATGAGACAAAGTCCGGGAGAACTTTCTCCG 792
Qy      416 yValSerAlaGlyGlnProGly---SerHisProSerProAlaProArgArgSer--- 433
Db      793 AAAAAGTGATGGGAGACCGCGGTGCGAGACGACCTC-CCAAAGACCAAGAGGGGCTGAG 851

```

```

Qy 434 -----ProbaGlyAsnYsgIn-AlaAspGlyAlaAlaSerAlaPro 447
Db 852 CAGCATCTGGATCCCGCTGGAACCGGGAGAGCCCGATTCCAGATTTTCAGACA 911
Qy 448 ThrGluGlu-GluGluGluValValYsAspArgThrGlnGlnSerYsProGlnProP 467
Db 912 CCTCAAGACAGAGGTGATGCGGAGAGCCAGGTGTCACAGCGGAAGGAGCGGG 971
Qy 467 ogInProSerThrGlySerLeuThrGlySerPro-----HisLe 480
Db 972 GATCCCTCTCGGAGCCGGAATATCGGGGCTCTGGCGAGACACCCCTGAGTCTCT 1031
Qy 480 userGluLeuSerValAsnSerGln----- 488
Db 1032 GCACCTCTGTCTCTGGACAGCCAGACCGGGAACAGGTCTACGCCCTCGAGACT 1091
Qy 489 -GlyGlyValAlaProAlaAsnValThrLeuSerProAsnLeuSerPro-----AspTh 506
Db 1092 GGGGGACTCATCCCGGAGGGCCCAATGAGCTGAGTCTTGCGGCCCCAGAGATAC 1151
Qy 506 rLyGlnAlaSerProLeuLeuSerProLeuLeuAsnArgGlnAlaCySProArgThrAs 526
Db 1152 CAGCAGAGGGGCGCAAAACGAGACCCCGAGCTGAGATGAGGGGAGCGGGCGGCTC 1211
Qy 526 P-----AspGluAspGluGlyArgArgGlyAspArgPheProThrArgLy 540
Db 1212 GGGAGCTGAGCTTACCAACAGAAAGACTCCCGGCTGGCGGAACTCTGCCCCAGAC 1271
Qy 540 AlaArgPheLeuAlaYs-----GluValSerTh 550
Db 1272 CTTGACAGCTGCTGCGCAAGCCAGGTGAGCGGAGAGGATCATCAGCAGCTGCTGGT 1331
Qy 550 rThGlnArgThrThyLeuYsAspLeuGluValIleTherSerThrPheGlnSerThrVa 570
Db 1332 GACAGAGCGGCGCCACGTCGTCATGCTGGGTCTGACACCTCTTCTCCAGCCCAT 1391
Qy 570 lSerLyGluAspAlaMetPro---GluAlaLeuYsSerLeuLeuPheProAsnPheG 589
Db 1392 GGCAGAAATGCTGTTCTTCCCTTGGAGAGACTCAGAAC---ATCTTCCCGAGCTGGA 1448
Qy 589 uProLeuNlYsYsPheNlYsThraPheLeuYsGlnlLeuGlnArgLeuAlaLeuTr 609
Db 1449 CAGACTATGAGGTGATCCCTGTTCTGATCGCTGATGAGCGG----- 1497
Qy 609 pGluGlyArgSerAsnAlaGlnIleArgAspTr-----GlnArglLeuYsAspValMe 627
Db 1498 -----AGCAGAGAGAGTGGCTATCTACAGAGAGATCGAGACGTGCT 1541
Qy 627 tLeuYsAsnIleGlnGlyMetYsNlYsLeuAlaAlaHisLeuTr----- 642
Db 1542 GCTGGCCCGGTTTATGCT-----GCTGAGGGCTCTCGTTCAGAAATCTC 1589
Qy 643 -----LysHisSerGluAlaLeuGlnAlaLeuGlnYsGlyYsE 657
Db 1590 CTCCTGCTTTCAGCGCCGAGTATGCTTTCAGACAGCTCAAGCCAAAGCCAA 1649
Qy 657 rSerArgArgLeuGlnYsAsnPheCySArgAspPheGluLeuGlnYsValCyS---TyrLe 676
Db 1650 GGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1709
Qy 676 uProLeuAsnThrPheLeuLeuArgProLeuNlYsArgLeuMetNlYsTrLyGlnValLe 696
Db 1710 GCAAGTAAAGACATGATCCCGAGAGATGACAGGCTGACCAAGTACCCCTGCTCT 1769
Qy 696 uGluArgLeuCySValYsNlYs---HisProSerHisAlaAspPheArgPhe--- 712
Db 1770 GCAAGACATCGGCGAGAAACAGAAAGCCAGAAACGAGAAATGAGTGGCGAGC 1829
Qy 713 -----CyArgGlnAlaLeuAlaGlnIleThrGlnMetValAlaGlnLeuNlYsGlyTh 730
Db 1830 CAGAGTGTGCGGGAAATTTTACACACGTCACCAAGCGCTGCTGATGAGAG---GA 1886

```

```

Qy 730 rMetIleYsMetGluAsnPheGlnYs-----LeuNlGluLeuYsYsAsp-- 746
Db 1887 CTGCTGAGGCTAAAGACATATACGCGGCTCGACTGTCCACCTTGGCAGACAG 1946
Qy 747 -----LeuIleGlyIleAspAsnLeuValValProGlyArgGluPheIleArgLe 763
Db 1947 GCACCTATCTGAGCGAGTTCAGAAACCTGACATACCAAGAAATTTGTTGCCAGA 2006
Qy 763 uGlySerLeuSer---LysLeuSerGlyYsGlyLeuGlnGlnArgMetPhePheLeuPh 782
Db 2007 GGGCCCACTGACTAGCGGGGTGACTAAGACAAAGCAGTGAAGGTGATGCTGCTGCT 2066
Qy 782 eAsnAspValLeuLeuYsTrThr-----Se 790
Db 2067 GAGACACTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2126
Qy 790 rArgGlyLeuThrAlaSerAsnGlnPheYsValHisGlyGlnLeuProLeu----- 807
Db 2127 CCGGACACTGACCGCCACCGCCGATGACAAAGACATGCTGCGGCCGTGCGGCTCAC 2186
Qy 808 -TyrGlyMetThrIleGluGluSerGluAspGluTrpGlyValProHisCySLeuThrLe 827
Db 2187 CTCGCCATATACCGGAGGTGCGCACCGATCACAAAGCTTCTACCTCTTTTACTCTG 2246
Qy 827 uArgGlyGlnArgGlnSerIleIleValAlaAlaSerSerArgSerGluMetGluYsTr 847
Db 2247 GAGACAGAGAGCCCAATATACGAGCTGTCGACAGACTGTGTCGAGCGGAAAACTG 2306
Qy 847 P-----ValGluAspIleGlnMetAlaIleAspLeuAlaGluYsSerSerSerPr 864
Db 2307 GTGTGCTCTCATCATGAGACTGCGGATCCCGAAAGTCCCTGCGCCCTCTGCGCC 2366
Qy 864 oAlaPro-----GluPheLeuAlaSerProProAspAs 876
Db 2367 TAAGCCCGGCGCCAGCGCCAGACAGACCCGAGAACCTTCTTACAGACTTGAAGACG 2426
Qy 876 nLySerProAspGluAlaThrAlaAlaAspGlnGluSerGluAspAspLeuSer 894
Db 2427 GAATGTGCGCGAGAGACGTCCTCAGCTGATGCGGAGCGAGAGAAATCTTCACT 2481

RESULT 5
US-60-796-903-635
; Sequence 635, Application US/60796903
; GENERAL INFORMATION:
; APPLICANT: Johnson, Patrick
; APPLICANT: Harkin, Paul
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Colorectal Cancer Microarrays and Methods of Use
; FILE REFERENCE: 55815-300 (327365)
; CURRENT APPLICATION NUMBER: US/60796,903
; CURRENT FILING DATE: 2006-05-02
; NUMBER OF SEQ ID NOS: 1071
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 635
; LENGTH: 4768
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-60-796-903-635

Alignment Scores:
Pred. No.: 0.0226 Length: 4768
Score: 174.00 Matches: 120
Percent Similarity: 36.9% Conservative: 96
Best Local Similarity: 20.5% Mismatches: 229
Query Match: 3.2% Indels: 143
DB: 8 Gaps: 27

US-09-555-342B-2 (1-1045) x US-60-796-903-635 (1-4768)
Qy 365 HisSerLySerLeuHisSerIleArgSerLeuAlaSerGluProThrGluLeuAsnSerGlu 384
Db 47 CATGAGACCTGGAAGACAGTGGCGGAGTGGCT---CATCATTGGCAA----- 91

```

```
Oy 385 ValLeuGlnGlnSerGlnGlnSerThrSerLeuThrPheGlyGlyGlyValAlaGlySerPro 404
Db 92 -----GGTGGTGGCCCAACCAACCG 112
Oy 405 GlyGlyGlnSerCysArgArgGlyLeuGlyProLeuValSerAlaGlyLeuProGlySer 424
Db 113 GGTGACCTGGGACTCGCTCAGGTGTTCGACCT-----TGGCGACAGCCCTCCGGGATGG 166
Oy 425 HisProSer---ProAlaProArgSerProAlaGlyAlaLeuValGlnAlaAspGly--- 442
Db 167 AGTCTGCTCTGCCAGCTGCTTAACAACCTCGGGCGACCTCCATCAACCTGAAGAGAT 226
Oy 443 -----AlaAlaSerAlaProThr-----GlnGlnGlnGlnGlnValValLeu 456
Db 227 CAACCTGAGCCCGCAGATGCTCCAGTTCTCTGTTGAAGAACATAGAGACATTTC--- 282
Oy 457 AspArgThrGlnGlnSerLeuProGlnProGlnProSerThrGlySerLeuThrGly 476
Db 283 -----TCACGGCTGTGTGAGACGTTTGAGATGAGAAAGTAACTTT 327
Oy 477 SerProHisLeuSerGlyLeuSerValAlaSerGlnGlyVal----- 491
Db 328 TCGAGGCACTTGACTTG-TTTGAATGTTCTGATCTTGGAAGGTTATAGAAACATTATCA 386
Oy 492 -----AlaProAlaAsnValThrLeuSerProAsnLeuSerPro----- 504
Db 387 CGACTTCTCGAACACCTATAGCATTTGGCCACAGGAATCAGGCCCTTCCCAACAGAA 446
Oy 505 -----AspThrLeuGlnAlaSerPro-----LeuLeuSerProLeu 517
Db 447 AGCATTAATGATGAAGACATCTACAAAGCCCTTGATTAATAGATGAAGAACCTTGTG 506
Oy 518 AsnAspGlnAla-----CysProArgThrAspAspGly----- 528
Db 507 GAAATGAAAGAAATCTGTATGACTGTGTTATAGGGAGATGAAGGTGAAAGTCTAT 566
Oy 529 -----AspGlnGlyArgArgLeuArgPhePro---ThrAspValAla 541
Db 567 GAGACCTTAATGAAGGACAGAGAGACACATCAGCCCAATGTCCGAAATGATTTACGA 626
Oy 542 TyrPheLeuAlaLeuGlyValAlaSerThrThrGlnArgThrTyrLeuValAspLeuGlyVal 561
Db 627 AGTTGTTGTCTAGCAGAAATTAAGCAGACAGAGAAATATATACGAAACTTTGGAGTCA 686
Oy 562 IleThrSerTrpPheGlnSerThrValSerLeuGlyAlaMetProGlnAlaLeuLys 581
Db 687 ATAGAAAAGTATTTTCATGCGACCACTAAAGAA-----TTTCTGACAGCAGCAGAAATT 740
Oy 582 SerLeuIlePheProAsnPheGluProLeuHisLeuPheHisGlnAspPheLeuGly 601
Db 741 GATTCACTATTCATCAACATTTCTGAACTTGTAAACTTTCATCGAACCTTAATGCAAGAG 800
Oy 602 IleGlnGlnArgLeuAlaLeuTrpGlnGlyArgSerAlaAlaGlnIleArgAspTyrGln 621
Db 801 ATTCATGAT----- 809
Oy 622 ArgIleGlyAspValMetLeuLysAsnIleGlnGlyMetLysHisLeu----- 637
Db 810 -----TCCATTGTAAATATAAAATGACCAAGAACTGTGACCAAGTTTATTATTAATAC 860
Oy 638 AlaAlaHisLeuTrpLysHisSerGlyAlaLeuGlnAlaLeuGlnAlaGlyIleLeuSer 657
Db 861 AAGGAAAGATTGGTTATTACGGGCACTACGACGTGAGAGTCAAGCCATCTCTAGT 920
Oy 658 SerArgArgLeuGlnAsnPheCysArgAspPheGlnLeuGln--LysValCysArgTyrLeu 676
Db 921 TTAGACTACATTTCTAAGTCAAAAGAGATGTCAAACTGAATATAGAGAAATGTTCCAAA 980
Oy 677 ProLeuAsnThr-----PheLeuLeuArg-----ProLeuHisArgLeu 689
Db 981 AGAGCAAAATTAATGGAAATTTACTCTTGAGACTTGCTTGCTTCAATGCAAGATGTT 1040
Oy 690 MetHisTyrLysGlnValLeuGlnArgLeuCysLysHis-----HisProProSerHis 707
```

```
Db 1041 TTAAGTACCAACCTTCTCCTCAGAGAACTGCTCAAACTACCACTGATCCGACTGAGAG 1100
Oy 708 AlaAspPheArgAspCysArgAlaAlaLeuAlaGlyIleThrGlnMetValAlaGlnLeu 727
Db 1101 GCAATCTGAAACCTGGCTCTTGATGCCATGAAAGACATTTGGCAATATGTAAATGAA--- 1157
Oy 728 HisGlyThrMetIleLysMetGluAsnPheGlnLysLeuHisGlyLeuLysAspLeu 747
Db 1158 -----GTAAGAAAGATTAAT--GAGACCTTGCTGAAATTAACAGTTTCAG 1202
Oy 748 IleGlyIleAspAsnLeuValAlaPro-----GlyArgGly 759
Db 1203 CTATCTATAGAGAAATTTGAACCAACGATTTTGCTTTTGGACGACCTCAAGGAGATGCT 1262
Oy 760 PheIleArgLeuGlySerLeuSerLysLeuSerGlyLysGlyLeuGlnArgMetPhe 779
Db 1263 GAAATTCGAATTAACCACTCTGACACAGCATACC-----AAACAAAGAAAGCATATTC 1313
Oy 780 PheLeuPheAsnAspValLeuLeu-----TyrThrSerArg 791
Db 1314 TTTCTTATTTGATTTGGCAGTATGCTATGTAAAGAAAGGTGATTAATGAAATGAAAG 1373
Oy 792 GlyLeuThrAlaSerAsnGlnPheLysValHisGlyGlnLeuProLeuTyrGlyMetThr 811
Db 1374 GAAATTAATAGTCTTTCAGCAGTACAGATACCAAT-----AATCCTTACA 1418
Oy 812 IleGlnGlnSerGluAspGlyTrpGlyValProHisCysLeuThrLeuArgGlyGlnArg 831
Db 1419 ACCCATTAAGAAACAAAAAGTGTCTTATGCTTCTTACCTCAATCCATACCAAGACAA 1478
Oy 832 GlnSerIleIleValAlaAlaSerSerArgSerGlnMetGlyLysTrpValGluAspIle 851
Db 1479 AATGGCTTAAGAAATTTATTTGCAAAACAAAGATTTAAAGAAATGGCTTAAGACAGTTT 1538
Oy 852 GlnMetAlaIleAspLeuAlaGlyLysSerSerSerProAlaProGluPheLeuAlaSer 871
Db 1539 GAATGGCTTTG-----TCTAACATTAAGACCAAGCTATGACAGCTCC 1580
Oy 872 SerProProAspAsnLys 877
Db 1581 AATTTCCAGCACTTCAAG 1598

RESULT 6
US-60-796-903-638
; Sequence 638, Application US/60796903
; GENERAL INFORMATION:
; APPLICANT: Johnston, Patrick
; APPLICANT: Harkins, Paul
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Colorectal Cancer Microarrays and Methods of Use
; FILE REFERENCE: 55815-300 (327365)
; CURRENT APPLICATION NUMBER: US/60796, 903
; CURRENT FILING DATE: 2006-05-02
; NUMBER OF SEQ ID NOS: 1071
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 638
; LENGTH: 13664
; TYPE: DNA
; ORGANISM: Homo Sapiens
; US-60-796-903-638

Alignment Scores:
Pred. No.: 1.29 Length: 13664
Score: 154.00 Matches: 205
Percent Similarity: 33.1% Conservative: 137
Best Local Similarity: 19.9% Mismatches: 347
Query Match: 2.8% Indels: 345
DB: Gaps: 48

US-09-555-342b-2 (1-1045) x US-60-796-903-638 (1-13664)
Oy 89 LysLysIleThrValTrpLeuAspLeuLeuAspPro-----IleValLysGlnIleArg 106
```


Db	1363	AGAAACCTCAGAGTTTAT-----CTGAAACCTTGAAATTACTTTTCAGAAAGG	14333
Oy	107	ArgProLyshLeVal-----ValValLyshPheValValLyshPhe	119
Db	1434	AGATCAAAAGAGAGTCAAGAAAGAAAGAAAGACACTTTTAAAAAGATGCCAACAGTA	14939
Oy	120	PheProPhePheThr-----GlnLeuGln	128
Db	1494	TTCTCTCCAGCAAGGACCATTAAGAAAGTCTGAGTTTGTTGATGACATACTACTAGAA	15533
Oy	129	GluGluLeuThrArgTyrLeuPheAlaLeuGlnValLyshGlnAspLeuAlaGlnLyArg	148
Db	1554	GAGAGCCCTT-----ATGATTCAGTTGAAGAGA-----	1580
Oy	149	LeuThrCyAsnAspThrSerAlaAlaLeuLeuLeuSerHisValGlnSerGluLe	168
Db	1561	-----GTAGTAAAGTGAATCACTCCTTCCACAAAGAA---	1616
Oy	169	GlyAspPheAspGluAlaLeuAspArgGlnHisLeuAlaTyAsnLyserTyrLe-----	186
Db	1617	-----AACGAAGAACTGATGTGTGGCTATCAAGCAGT-GATGAGATATAACAGSAA	1666
Oy	187	-----ProGlnAspAlaLeuGluAspLyshValGlnPheHis	200
Db	1667	GACAGCCAAAGATCAGGGGAAAGAGAGAGCACTTCATAAAACAATCATGAATGAATGAGT	1726
Oy	201	HisAsnHisLeuGlyGlnThrProAlaGluSerAspPhe-----GlnLeuLeuGln	217
Db	1727	GCTGTGAAGATGCTTCACAGCTTCTGAAAGATGATGATGATGATTCATCAGAAACCACTCAAGAG	1786
Oy	218	IleAlaArgArgLeuGluMetTyrGlyIleArgLeuHisProAlaTyAspArgGluGly	237
Db	1787	ATTAGTACAGACTGACTGAG-----AGCCAAAGAGAGAGAA	1822
Oy	238	ThrLyshIleAsnLeuAlaValAlaAsnThrGlyIleLeuValPheGlnGlyPheThrLysh	257
Db	1833	ACAAAG-----GGAAAAAGCCAAA	1840
Oy	258	IleAsnAlaPheAsnThrAlaValArgLyshLeuSerPheLyshArgLysh-ArgPheLe	277
Db	1841	ATTAACAGCA-----GGGAAACACACACGACTCTGAAAAAGTACACAAAGCACTT	1891
Oy	277	ValLeuLyshArgProAspAlaAsnSerAlaTyrGlnAspThrLeuGlnPheLeuMetAl	297
Db	1892	GATGAAGATGCAGAAAGACGTCACTCAAGCGCATGATGAAGACGATGAGCACTTGATGA	1951
Oy	297	AspArgAspPheCyshSerPheTyrLyshCysValGlnHisAlaPheAsp	317
Db	1952	AGTCTGAACTT-----	1963
Oy	317	GluPheGluGluProLyshProLyshProLyshProValLeuPheSerArgLyshSerSerPh	337
Db	1964	AAATATCAGAACTAAAGCTGAGAAAGTGAAGAACTTTAGTTACTGAGAA-----	2017
Oy	337	AspArgPheSerGlyArgThrGlnLyshGlnValLeuAspTyrValLyshGluGlyGlyHisLysh	357
Db	2018	-----GAGGGCTACG	2028
Oy	357	AluValGlnPheGluArgLyshAspSerLyshHisSerIleArgSerIleuAlaSerGln	377
Db	2029	CGGATTT-----AA	2037
Oy	377	AspThrGluLeuAsnSerGluVal-----LeuGluGlnSerGlnGlnInse	392
Db	2038	AACAATTAGCTCAACAGTACATGACATGATATAATTTCTCGAGAGTCATCACAAGAAA	2097
Oy	392	ThrSerLeuThrPhe-----	397
Db	2098	AACAAAGTTGTAATTTTGACGAAGACCCGAATTGGAATGGAAACCTCAACAGACTACC	2157
Oy	398	-----GlyGluGlyAlaGluSerProGlyGluGlnSerCyAspArgGln	412

Db	2158	TGAAGATAGGTCAAGGGGAGAGGGATCTTCAGTCTGCATGCTTCACGCTTCACCTCCGG	221
Qy	412	YLYSGIuPro---LYsValSerAlaGIuGIuProGIySerhIaProSerProAlaProAr	431
Db	2218	TACATCCCTTCATCAGATGATCATCATCTGATGAGAGACAGTACAGTAGCCGAGTACAA	2277
Qy	431	gArSerProAlaGIuAsnLYsGIu-----	439
Db	2278	AAAAGAGAGGCAACACGCAACCCAAAGCTCGCACAGACACATGGCCCTCTTTGGC	2331
Qy	440	-----AlaAspGIuAlaIaSerAlaProThrGIuGIuGIuValIaLYsAs	457
Db	2338	TACTATTGAGATTTCTTCAGAGGAAGAAATTGAGAGAGGAAGAAAGATTATTAAAGCA	2399
Qy	457	PaGIThrGIuInSerLYsProGIuProGIuProSerThrGIySerLeuThrGIySe	477
Db	2398	GCAAGAAAGACAGAGGGGAAATAGAACGACAAACAAAGAAAGGTTCTGTAATAAATCAAA	2455
Qy	477	rProhIeLeuSerGIuLeuSerVal-----AsnSerGIuGIyValIaProAl	494
Db	2458	GAAAGACAAAGATGAACTTCAGAGCTCAGAGAGAAAGGAGGAGCCCAAGACACCACTTGA	2511
Qy	494	aaAsValThrLeuSerProAsnLeuSerPro-----AspThrLYsGIuAlaSerProLe	512
Db	2518	TAACTCTCTCCATTGAAAGTGCATCTCCGACAAAGAGTTAGCTCAGGCTCGCAGAAT	2577
Qy	512	uIeSerProLeuLeuAsnArgInAlaCYs-----ProArgThrAspAspGI	528
Db	2578	GGAGAG-----CTCCATAGATCTCTGTTCTTGAAATATTCACCTAGACATAGATGCA	2631
Qy	528	uAspGIuGIy-----ArgArGIyAsArgPheProThrAspLYsAlaTYrPheI	544
Db	2632	CCCAAGAGGTTTGAAATTAAGCCCGGAAATAAATATGAAAGTACAAAGTTTAAATTT	2691
Qy	544	eAlaLYsGIuValSerThrThrGIuAGThrTYrLeuLYsAspLeuGIuValIeThSe	564
Db	2692	GCCCAACAGCTGTTCATTATTAATCTCACCAAC-----	2722
Qy	564	rTrpPheInSerThrValSerLYsGIuAspAlaMetProGIuAlaLeuLYsSerLeuI	584
Db	2723	----GATGAGCAATCTTATTAAGCAAGAAAGATGAGCCAAAGGCGTTAAAAAGTCTGA	2778
Qy	584	ePheProAsnPheGIuProLeuNHLYsPheHis-----ThrAs	597
Db	2779	GGAAGATGATGAAGAAATGATGATTAACACACAAATATCAAAAGCTTTCCAGCTGCAAA	2838
Qy	597	nPheLeuLYsGIuIeGIuGIuArgLeuAlaLeuTrpGIuGIyArgSerAsnIaGIuI	617
Db	2839	TGAACGAGATGAAAGTGTGAAAAAGAGCCTTTGTATGATGGG-----ATGCTAAT	2889
Qy	617	eArgAspTYr-----GlnArgIeGIyAspValMetLeuLYsAsnIeGIuGIyMe	634
Db	2880	AGAGGATTTATTATTATGAAATTTTAGTAGAAGACACGTACAAATGATGCGTAGATGCG--	2944
Qy	634	LYsHisLeuAlaIaHisLeuTrpLYsHisSerGIuAlaLeuGIuValaLeuGIuAsnGI	654
Db	2948	-----AGTCTGCTTAACAGGCAAGAAAGAA-----GAAATAGG	2979
Qy	654	YLYsSerSerArg-----ArgLeuGIuAsnPheCYsArgAspPh	668
Db	2980	ATTATATGACAGCAAGAAAGAGACAGCAAAAGATTAAGACTTTTCAGAACAGATTTATGAAG	3039
Qy	668	eGIuLeuGIuLYsValCYsTYrLeuProLeuAsnThrPheLeuLeuArgProLeuNHAsr	688
Db	3040	TCCTATATGAGAAATATACAGACCTCCGAAAGAGAGTTTATAGTTTGAAGAGCTTACATTC	3099
Qy	688	gLeuMetNHsTYrLYsGIuValLeuGIuArgLeuCYsLYsHisIAsrProSerSerhIaI	708
Db	3100	TGTTGTGCGCTCAGGAAGATATTGTTCAAG---TCTTTTATCATCCAGAAAGCAT--	3154
Qy	708	aAspPheArgAspCYsArgAlaIaLeuAlaGIuIeThrGIuMetCYsAlaGIuLeuH	728
Db	3155	-----GAGATGTGGAACCTG-----	3169

```

Oy      728  eGlyThrmetlle---LyMetGluaBPhelInlySleu-----              740
Db      3170 -GGTAAGTATGTGAACCTTTCAACGAAGAAGAAAGAACTACTAAGTGCTGATGCTGCCTA 3228
Oy      741  -HieGluLeuYlVtVAePleu-----                          747
Db      3229 TGMAGAACCTTTATGMAAGAGCACAAGATGCATTAAACCCTGGATGTAAAGCCCAAGCCAGGC 3286
Oy      748  ---IleGlyIleaPeaNeUvaIvalProGIYAArgLupheIeaArgLeu-----GI 764
Db      3289 CCCCATTGGT---GAGGATATGACAGAGTCACCATGGACCTTGACAGAAATGCCAGATGC 3345
Oy      764  ySerLeu-----SerYlVleSeSerGilyVyleGlyLeuGlnGlnaGmetPhePheLeuPh 782
Db      3346 CTCTTTTGACATCAAAAGTGTCTCTCCAGGAGCGCTCTTCAAGATTGC----- 3391
Oy      782  eAsnaBvValleuLeuyrThrSerArGilyLeuThraIaseRaenGlnPheYValHi 802
Db      3392 -----ACGACGACGTGCAACACTCTCATCC 3417
Oy      802  eGlyGlnLeuProLeuTyrgIymeTrhlIeGlnGluSeSerGluAepGIutRgilyValPr 822
Db      3418 AGATGTATAAATTAACCCAACATTTTTCACAGAGAAGAAATTAGATGATATGTAAACCGA 3477
Oy      822  oHieCybeYleuthrLeuArGily-----GlnArGlnImberIleIleValAlaIalaSeSe 840
Db      3478 TTATTCACAGAGAAATTCAGAGATTAATGCCCATTAATCGCGATTTTGACCTACTCGGA 3537
Oy      840  rArSerGIumeGilyStRpVaIGluApilIeGlnMeAlaIleaPleuAlaGluLy 860
Db      3538 GCCTTCAGAAAGTGTACATCTGTC-----CCACC 3566
Oy      860  sSeSerSerProAlaProGIuPheLeuAlaSeSerProBpaBaulySeSerProAs 880
Db      3568 CTCGACACACACTTCTCTCACATCATCTGTGTTCTTCGGTCTGTACCAAGATAGCTTTC 3627
Oy      880  pGluAlaPhrAlaAlaAepGln----- 887
Db      3628 ACCCATTAATCCCTGATAGCAATAACCAACAGTTTATACAGAGCCAGTGCATGANTAAC 3687
Oy      888  ----GluSeSerGluAeAPleuSeSerAlaser----- 896
Db      3688 TAAATTGAAAGATTCTGAGAAATTTCTTCATCAACTTATTTCACAGCAGATTATAGA 3747
Oy      897  -----ArgThrSeSerLeuGlnArgGlnAlaProhibArGilyaBnThme 911
Db      3748 CTATCCAGAAAGAAATTAAGTATCTTTAGATCGAGCTCCCCACAGATGATAGAGCTAG 3807
Oy      911  tVal---HisValCystrPhIseArGaBnThrSeSerValserMeVeLaAPheserIleAl 930
Db      3808 TGCGATCATATGTATTATTCCTTATTCGATATAGCAATCTTCTCATATGAATCTGTAAGT 3867
Oy      930  a-----ValGluAenGlnLeuSeSerGlyAenLeu----- 940
Db      3868 ACCTAAACCTGAAGGCCAGCTTCTGACAGACGTGTTCTTCACTGACTTACTTAAATCTGAABA 3927
Oy      941  -----ArgLyPheLyBAenSeArBngily 948
Db      3928 GGACCCAGTGAAGAAAGCCAGAAAGAAAGAAATCGG 3961

RESULT 7
US-10-529-655-22
; Sequence 22, Application US/10529655
; GENERAL INFORMATION:
; APPLICANT: Ludwig Institute for Cancer Research, et al.
; TITLE OF INVENTION: Human Sarcoma-Associated Antigens
; FILE REFERENCE: L0461.70156M000
; CURRENT APPLICATION NUMBER: US/10/529,655
; CURRENT FILING DATE: 2005-03-30
; NUMBER OF SEQ ID NOS: 167
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 22
```

[illegible]

Db 2098 -----CTGCAACAAAGCCAGCTTGACAACTTCACAGCTCAGAGGTACT--- 2142
 Qy 240 IleasleuAlaValAlaAenThrGlyIleleuValPheGlnGlyPheThrIleAsn 259
 Db 2143 ----TCTCAGACTCAGCAACAGTAGCTCTGCTTATCTTCAGAGTCTTATTTCGACG 2199
 Qy 260 AlaPheanThrAlaValAlaValArgLysLeuSerPheLysArgLysArgPheLeuIleLys 279
 Db 2200 CAACCTGAATATACAGAAAGCTCCAGAGGTTGTACGACAGAAAGAGCCCAACTG----- 2253
 Qy 280 LeuArgProAspAlaAenSerAlaTyrGlnAspThrLeuGlnPheLeuMetAlaSerArg 299
 Db 2254 -----CCTGATGCTCAACAAATGTGTGCAATTTGTAAAGCTGACGACACAGAGAGTGA 2307
 Qy 300 AspPheCysLysSerPheThrPheIleCysValGlu---HisIleAlaPhePheArgLeu 318
 Db 2308 CAGCAAGAAAGAGCTTCTTGAAACATACACAGAAATTGCGAAAGCTTCGACAGACTA 2367
 Qy 319 PheGlnGluProLysProLysProLysProValLeuPheSerArgLysSerSerPheArg 338
 Db 2368 CAAGAAAGAAATTGACGAATTAAGAGCCAAAGCTGCTGCGAGGCTGAAATAATCAAT 2427
 Qy 339 PheSerGlyArgThrGlnLysGlnValLeuAspTyrValLysGlnGlyHisLysLys 358
 Db 2428 ---GAGATTGAAAGCCAGAAACAAACATCCAGACCTAAACCATGCTTGAGTCACAG 2484
 Qy 359 ValGlnPheGlnArgLysIleSerLysIleHisSerIleArgSerLeuAlaSerGlnPro 378
 Db 2485 GAGCAGTTGCTTCAGAAATTCGGAGCTCTTACAGATATGAGTAATCTCAGACAAACC 2544
 Qy 379 ThrGlnLeuAsnSerGlnValLeuGlnGlnSerGlnSer 392
 Db 2545 CTTGAAACAAATGAATGTTGCTTGAGAAACTTCGACGACCAATACATGATTAAGCTGT 2604
 Qy 393 -----TherSerLeuThrPheGlnGlyGlnValGlu 402
 Db 2605 GCTCTGAGCGGGCTATAGATGAAATAATCTCTCTAGAGAGAAAGAAA-AGAACT 2663
 Qy 403 SerProGlyGlnSerCysArgArgGlyLysGlu----- 414
 Db 2664 GCGCAGACTTCCTTCTGCTGTGAGAGCCAGATCAGACTTGAGAGACTCGCCGATGT 2723
 Qy 415 ProLysVal-----SerAlaGlyGlnProGlySerHisProSerPro 428
 Db 2724 CCTCTCTCCAAATGAACTACTATGACAAAGTATGAGAG----- 2762
 Qy 429 AlaProArgArgSerProAlaGlyAsnLysGlnAlaAspGlyAlaIle----- 444
 Db 2763 -----TCTCTAGAGGGCCAAAGGCTGGAAGTGGAACAGTTATCTACTACTG 2810
 Qy 445 SerAlaProThrGlnGlnGlu-GlnGluValValLysAspArgThrGlnGlnSerLysPr 464
 Db 2811 TCAAAACTCCAGTGGGTGAAGAAGAAATG-----GAAACCAATTTAGCCCTTG 2861
 Qy 464 GlnProProGlnProSer---ThrGlySerLeuThrGlySerProHisLeuSerGluLe 483
 Db 2862 GCAGAAAGAAACAAGAGATCATTCAGCAGTTACAGACGCTCTCTTAT----- 2910
 Qy 483 userValAsnSerGlnGlyValAlaProAlaAsnValThrLeuSerProAsnLeuSe 503
 Db 2911 -----GATAGGAACAAAGAGTGAAGATCTTAAGCAACACTGCTGCAAACTTGG 2963
 Qy 503 rProAspThrLysGlnAlaSerProLeuLysSerProLeuLeuAsnAspGlnAlaCysPr 523
 Db 2964 ACCAGGCGCAGAGTAG-----ATAGCAAGAGAGCTGTCCCA 2999
 Qy 523 oArgThrAspAspGlnAspGlnGlyArgArgLysArgPheProThrAspLysAlaTyrPh 543
 Db 3000 GCGGTCTACAG-----CGAAAGAAAGATGCTGACGAGACTT----- 3036
 Qy 543 eIleAlaLysGlnValSerThrThrGlnArgThrThrLysLeuAspLeuGlnValIleTh 563
 Db 3037 -----CTAAGTATCGAAATTAACAAAGTCTGGAACATGAATGAGATTCA 3083

Qy 563 rSerTrpPheGlnSerThrValSerLysGluAspAlaMetProGlnAlaLeuLysSerLe 583
 Db 3084 AGGCGCTCTCACTCTGTGAGACACAGAGGACAGAAAGCAAGCTCTGCAGAGAGTT 3143
 Qy 583 uIlePheProAsnPheGlnProLeuHisLysPheHisThrAsnPheLeuLysGluIleG 603
 Db 3144 GGTGCAAGCCCTTATATGAA-----AGAAATTCAGAAATTAAGGCCCTCGCG 3188
 Qy 603 uGlnArgLeuAlaLeuTrpGlnGlyArg-----SerAsnAlaGlnIleArgAs 619
 Db 3189 CCAATATTATGGA-----GGAGAGACTCCCTGATGTCCCAAGACCCATCTTAA 3239
 Qy 619 pTyrGln-----ArgIleGlyAspValMetLeuLysAsnIleG 632
 Db 3240 CCAACAAAGCTGAAGTTAACCCCACTGGCGCTTTGAAA-----CAGACTGATCA 3290
 Qy 632 nGlyMetLysHisLeuAlaAlaHisLeuTrpLysHisSerGlnAlaLeuGlnAlaLeuG 652
 Db 3291 AGGTCAATGCAAGATACCTTCCAGA-----GATGATAGCACTTCATTTGACCAAGA 3344
 Qy 652 uAenGlyIleLysSerSerArgArgLeuGlnAsnPheCysArgAspPheGlnLeuGlnL 672
 Db 3345 GGAATGACAGATCCAGATCCACATTAAGGAGCTTGACACAGCTTGACAGGCTGGA 3404
 Qy 672 sValCysTyrLeuProLeuAsnThrPheLeuLeuArgProLeuHisArgLeuMetHisT 692
 Db 3405 A-----GAACTGAGTAAATGC 3419
 Qy 692 rLysGlnValLeuGlnLysLeuCysLysHisHisProProSerHisAlaAspPheArgAs 712
 Db 3420 CAAAGAGAACTTGAACTCATGCTGCTAAAGAA-----AAGA 3458
 Qy 712 pCysArgAlaAlaLeuAlaGlnIleThrGlnMetValAla-----GlnLeuHis 728
 Db 3459 AAGTCAGATGAAACTTGTGCTCTTACAGTCCATGATGCTGTGACAGAAAGAGAGCTGCA 3518
 Qy 728 sGlyThrMetIleLysMetGlnAsnPheGlnLysLeuHisGlnLeuLysAspLeuI 748
 Db 3519 GGTGAGAGGCTGTGATATGAGTGTCTGACACAGAAACATGACATTAAGAAATCTCAT 3578
 Qy 748 eGly-----IleAsnLeuValValProGlyArgGlnPheIleAr 762
 Db 3579 AAAGAACCTGCAATGCAACTGTGATCTCGAAGACATACCAAGCTATGGAA----- 3630
 Qy 762 gLeuGlySerLeuSerLysLeuSerGlyLysGlnGlnArgMetPhePheLeuPh 782
 Db 3631 -----GCCCTGACCCAGAAAGTCTTACTTCTTG 3659
 Qy 782 eAsnAspValLeuLeuTyrThrSerArgLysLeuThrAlaSerAsnGlnPheLysValHi 802
 Db 3660 GGAAGAAAGTTGCTTCAGTAGAATCCAGAGGTCAAGAAATTTACAGAAACCGAAGACAA 3719
 Qy 802 sGlyGlnLeuProLeuTyrGlyMetThrIleGlnGlnSerGlnAspGlnTrpGlyValPr 822
 Db 3720 GTTGCTGTGATGCTGAAGAGACTAGATGAACGAGTGG-----CTCAA 3767
 Qy 822 oHisCysLeuThrLeuArgGlyGlnArgGlnSerIleIleValAlaAlaSerSerArgSe 842
 Db 3768 TGAAGCTTAAAGACAGAGACAGAGCTCTATAGCACTGTGTGAATGCCATCCATCC 3827
 Qy 842 rGlnMetGlnLysTrpValGlnAspIleGlnMetAlaIleAspLeuAlaGlnLysSerSe 862
 Db 3828 AGAGAGCTCTGAAGAGACCGAACTGTGCGAGGTGAAGCTGAGGCTCAGGTGTACG 3887
 Qy 862 rSerProAlaProGlnPheLeuAlaSerProPro-----AspAsnLys 877
 Db 3888 CAGTTCGCTGAAGAAAGTTCTTGGAAGAGCTTGAGCCGCTTAACAGAGCTGAGACCT 3947
 Qy 877 sSerProAspGlnAlaThrAlaAlaAspGlnGlnSerGlnAspAspLeuSerAlaSerAr 897
 Db 3948 GCGCCGCAATTGAGGTGACGCTGACGGGATGACACCGAAGATACAAAGCACTGAGTTCA 4007

Qy 897 gThrSerLeuGluArgGlnAlaProHisArgGlyAsn 909
Db 4008 TGACAGTATTGAGAGAGGCTGCACACCATATGTCAC 4044

RESULT 8
US-10-576-496-19
Sequence 19, Application US/10576496
GENERAL INFORMATION:
APPLICANT: Teijin Pharma Limited
APPLICANT: Komori, Toshihisa
APPLICANT: Kanetani, Naoko
APPLICANT: Yoshida, Caroline Andrea
APPLICANT: Zama, Akira
APPLICANT: Kobayashi, Shinji
APPLICANT: Yamana, Kei
TITLE OF INVENTION: BONE AND/OR JOINT-DISEASE-ASSOCIATES GENES
FILE REFERENCE: Q94468
CURRENT APPLICATION NUMBER: US/10/576,496
PRIOR FILING DATE: 2006-04-20
PRIOR APPLICATION NUMBER: PCT/JP2004/015879
PRIOR FILING DATE: 2004-10-20
PRIOR APPLICATION NUMBER: JP 2003-359172
PRIOR FILING DATE: 2003-10-20
NUMBER OF SEQ ID NOS: 114
SOFTWARE: PatentIn version 3.1
SEQ ID NO 19
LENGTH: 3276
TYPE: DNA
ORGANISM: Mouse
FEATURE:
NAME/KEY: CDS
LOCATION: (114)..(2763)
US-10-576-496-19

Alignment Scores:
Pred. No.: 5.35 Length: 3276
Score: 131.50 Matches: 122
Percent Similarity: 34.7% Conservative: 92
Best Local Similarity: 19.8% Mismatches: 197
Query Match: 2.4% Indels: 206
DB: Gaps: 30

US-09-555-342b-2 (1-1045) x US-10-576-496-19 (1-3276)

Qy 453 GluValValIysAspArgThrGlnGlnSerIysProGlnProGlnProSer----- 470
Db 43 GAAAGTGTGAGGACACAGAGCGACCTGCATACCTGAGCCTCCAGAGAGAAAGCTGCAC 102
Qy 471 -----ThrGlySerLeuThr-----GlySer 477
Db 103 TCTGGCATCCAGTACAGACAGAGACTAAGCATAAAGAGCCATGACTTCTGCGCCT 162
Qy 478 ProHisLeuSerGluLeuSerValAsnSerGlnGlyValAlaPro----- 493
Db 163 AGTCATATGCTGGAAGCTGCCCTGGAGAGATGACGGGATCATTCAGAGCACTAAACA 222
Qy 494 -----AlaAsnValThrIleuSerProAsnLeuSerProAspThrIys----- 507
Db 223 GCTGCAGATTTTATGATGATGTAACCTGTGAGCCTGGCTGTCTCCCATTCACCTGCTTG 282
Qy 508 GlnAlaSerProLeuIleSerProLeuLeuAsnAsp-----Gln 520
Db 283 AACTTCATGCTGTGCTCAT-----CTCATTAAGAGCCCTGAGACATGACCTTGAGATGTTG 339
Qy 521 AlaCybProArgThrAspAspGluAspGluArgArgIysArgPheProThrAspIys 540
Db 340 GCACCTCTCTCAG-----GAAAGAGAGAGCCCTCGAGCAGAGTCCCTGGCCCAACA 390
Qy 541 AlaTyrPheIleAlaIysGluValSerThrThrGluArgThrTyrIleuIysAspLeuGlu 560
Db 391 GCT-----ACCTACATTAAGAG----- 408
Qy 561 ValIleThrSerTrpPheGlnSerThrValSerIysGluAspAlaMetProGlnAlaLeu 580

Db 409 -----TGTGTTGAGGACAGCTTGCCAGGTAATAC----- 441
Qy 581 LysSerLeuIlePheProAsnPheGluProLeuHisIysPheHisThrAsnPheLeuIys 600
Db 442 -----CACGGTCTGCTAGTAAT-----GAA 462
Qy 601 GluIleGluGlnArgLeuAlaLeuTrpGluIysArgSerAsnAlaGlnIleArgAspTyr 620
Db 463 ACCTAACGAGAACCCCTTGCCACGGCTAGAGAGATAAAGAGTCCCTCATCTACAGGTG 522
Qy 621 GlnArgIleGlyAspValMetLeuIysAsnIleGlnGlyMetIys----- 635
Db 523 AGCCCTCCACAGATCAAGTG-----GAAGCCCAAGAGAGAGATTCGGACCTTGAA 576
Qy 636 ---HisLeuAlaAlaHisIleTrpIysHisSerGluAlaLeuGluAlaLeuGluAsnGly 654
Db 577 GTGTGTCTGGAAGGCCACACAGGTGAAGCTCAATCTGCAGAAAGATGCTTCAGCAGAG 636
Qy 655 IleYsSerSerArgArgLeuGluAsnPheCybArgAspPheGluLeuGlnIysValCys 674
Db 637 CTGCTAACTCGCAACATCTTGGAAGCCAGAAAGTTGATCTG---ATGACTGAGGTCT 693
Qy 675 TyrLeuProLeuAsnThrPheLeuLeuArgPro----- 685
Db 694 GAGCTGAAGCTCAAGCTGTTGTATGGAAGAAACAGAAAGAACAGAAAGAAACAG 753
Qy 686 -----LeuHisArgLeuMetHisTyrIysGlnValLeuGluArgLeu 699
Db 754 AGAAAGAGAGAGAGATTACTGCAAGAGCTTAAAGCACTCAAAATCAAGGTAGAGAGCTG 813
Qy 700 CysIysHisHisProProSerHisAlaAspPheArgAspArgCybArgAlaAlaLeuAlaGlu 719
Db 814 GAGAACGAGCGGAACCAATATGATGAGTGGAGCTGAAGGCCACCAAGCTGAGGTAGCCAG 873
Qy 720 IleThrGluMetValAla-----GlnLeuHisGlyThrMetIle 732
Db 874 CTGCAGAAACAGAGTGGGCCCTTAAAGATGACAGAAATTGAGCGCTCCACAGCAGCTGCC 933
Qy 733 Lys-----MetGluAsnPheGlnIysLeuHisGluLeuIys 744
Db 934 CGGAGTGCAGCTTCCACAGGACCATGACAGAGAGATCAAGAAATCCACCGTCTGAAA 993
Qy 745 LysAspLeuIleGlyIleAspAsnLeuValProGlyArgGlu---PheIleArgLeu 763
Db 994 -----ArgGAGATGGAACACTGCTGTGTCATAGAGTAAGACCGTGGATA 1044
Qy 764 GlySerLeuSerIysLeuSerGlyIysGlyLeuGlnArgMetPhePheLeuPheAsn 783
Db 1045 GAGGAGCTGACAGGCGCTGTTGAACAATTAAGGGTAAAG----- 1086
Qy 784 AspValLeuLeuTyrThr-----SerArgIysLeuThrAlaSerAsnGlnPheIys 800
Db 1087 GAGATTGTGATGGCAACTCAGAGGCGCTTCAGAAAGAACCCCTCATTCATGAAGATGAA 1146
Qy 801 ValHisGlyGlnLeuProLeuTyrGlyMetThrIleGlnGluSerGluAspGluTrpGly 820
Db 1147 ATTAAGGGAAGCTTCCGAAATGGAATGCCAATTAAGACCCAGAGAAA----- 1197
Qy 821 ValProHisCysLeuThrLeuArgGlyGlnArgGlnSerIleIleValAlaAlaSerSer 840
Db 1198 GTCCCG-----AAGCA----- 1209
Qy 841 ArgSerGluMetGluIysTrpValGluAspIleGlnMetAlaIleAspLeuAlaGluIys 860
Db 1210 -----GAGATTCACACAGCA 1224
Qy 861 SerSerSerProAlaProGluPheLeuAlaSerSerProProAspAsnIysSerProAsp 880
Db 1225 TGCAGCTTCCACCCCA-----GAGCCACTT-----CCTTTG 1257
Qy 881 GluAlaThrAlaAlaAspGlnGluSerGluAspAspLeuSerAlaSerArgThrSerLeu 900

```

Db      1258 CCAAGAAATCATCTGAGAGCAGGCTCAGAGAAACTCTTCAGTCTAGAAAGACTTG 1317
Qy      901 GUAArglnAlaProHisAryS-----GlyAanHmetValHisValCysTrpHis 917
Db      1318 AGAGGTGAATCTGGGAGTAAATGTGTGATGGAAACAGCTGTCCCACTGGAGAGCC 1377
Qy      918 ArgAanHrSerValSerMetValAspPheSerIleAlaValGluAanGlnLeuSerGly 937
Db      1378 AAGACAGACTCTTCC-----CTACGGAGCAGAAATACCCC 1413
Qy      938 AsnLeu-LeuArgLysPheIlySerAenSerAenGlyTrpGlnLysLeuTrpValValPhe 957
Db      1414 ACATTACTGGAGACTTT-----CAGAGGCACACCAATGAGAAAGCTGCC 1461
Qy      957 rAanPheCysLeuPhePheTyrlYsSerHieGlnAspAanHisProLeuAlaSerLeuPr 977
Db      1433 -----CAGAGGCACACCAATGAGAAAGCTGCC 1461
Qy      977 OLeuLeuGlyTyrlSerLeuThrIleProSerGlnSerGlnAlaAnIleGln 993
Db      1462 AAATCTC-----CTCCCACTGCCTCCCTCAGGCTGACTCTTG 1501

```

RESULT 9

```

US-10-934-272A-1
; Sequence 1, Application US/10934272A
; GENERAL INFORMATION:
; APPLICANT: Rosen, Steven D.
; APPLICANT: Sasetti, Christopher M.
; TITLE OF INVENTION: Selectin Ligands
; FILE REFERENCE: UCAL-097CON2
; CURRENT APPLICATION NUMBER: US/10/934, 272A
; PRIOR FILING DATE: 2004-09-03
; PRIOR APPLICATION NUMBER: 09/860,298
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: US 09/243,560
; PRIOR FILING DATE: 1999-02-03
; PRIOR APPLICATION NUMBER: US 60/074,389
; PRIOR FILING DATE: 1998-02-11
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2269
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-934-272A-1

```

Alignment Scores:

```

Pred. No.: 3.69 Length: 2269
Score: 131.00 Matches: 135
Percent Similarity: 32.3% Conservative: 68
Best Local Similarity: 21.5% Mismatches: 211
Query Match: 2.4% Indels: 216
DB: 6 Gaps: 29

```

US-09-555-342B-2 (1-1045) x US-10-934-272A-1 (1-2269)

```

Qy      5 GlnGlnArgPro-ThrProGlySerArgLeuGlyAlaProGluAnSerGly----- 21
Db      647 CAGCAGAGCCAGACCCACAGGGGCCACCAAAAGCAGCATGAAAGACTCCGGGACCAAGGC 706
Qy      22 -IleSerThrLeuGlnArgLysGlnLysProProThr-----ProSerGlyLys 38
Db      707 CTCATCAGTGTGAGGTGAGAGACACAGCATGGGCGCCAGCTGTGCTGCTTCAAGTAC 766
Qy      38 eLeuValSerIleLysIleGlnMetLeuAspPheThr-----GlnGlnAla----- 53
Db      767 CCCAACTACAGTGCCTCCGGGGACACAGACTCCACCAAGCCAAAGCAGAGGCCACAGT 826
Qy      54 -----PheGlnValProGlnArgAlaProGlyLysVa 64
Db      827 GCTGCCAGCTGCAGGGCTTGGGGGTAGAGTTGAGGCTCTTCAGGAACCAAGCAAGAAAGC 886
Qy      64 lleuLeuAspAlaValCysAanHisLeuAanLeuValGlnGlyAanPheTr----- 80

```

```

Db      887 CACTGCAGAGCAGACTGT-----TTGTCTGGCCAGCAGCAGAGAGTGC 931
Qy      81 -----PheGlyLeuGlnPheProAanHisLys 89
Db      932 GGCCTTGCTTCAATCCCTTCAAAACACAGCTCCCAAGTGGGGCCGAGACCAAGTGAAGA 991
Qy      89 slyIleThrValTrpLeuAspLeuLeuLysProIleValLysGlnIleArgTrpLys 109
Db      992 TCCCTTGCTGCTAGAACTCAGCCCTTCCCACTGGCCCTGGA----- 1037
Qy      109 sHisValValLysPheValValLysPheProProAanHisThrGlnLeuGlnI 129
Db      1038 -----GACATGGAACAGACCTTCTCTGCTACCTTGGGAGCAAGA 1078
Qy      129 uGlnLeuThrArgTyrlLeuPheAlaLeuGlnValLysGlnAspLeuAlaGlnIlyArgLe 149
Db      1079 AGATCTACACAGACCTCTTAGAAGGCGAGCACTGA-----GCTCAATCCAGAT 1132
Qy      149 u-----ThrCysAanAspThrSerAlaAlaLeuLeuIleSerH 162
Db      1133 ACCCTGGAGTTCACGAGGTGATCTGCAGAGACTGGAGCAATCTGCTGGGAAAACTA 1192
Qy      162 sIleValGlnSerGlnIleGlyAspPheArgLysAlaLeuAspArgLysIleValAla 182
Db      1193 CATCATCTGACATGACAGAGAAACATAGACTGTGAGTGTTCGGCAGACCGGGGG-- 1250
Qy      182 sAanLysTyrlIleProGlnGlnAspAlaLeuGlnAspLysIleVal-----G 198
Db      1251 -----CACAGCTCTGTGGCTGGCGTGGAGAAAGTGTCTCCGCCCATGGCAG 1297
Qy      198 uPheHisHisAan-----HisIleGly--GlnThrProAlaGlnSerAspPheGlnLe 215
Db      1298 TGGCAGACCAATGGGGCCCTGGCAGACATCTCTGAGCAAGCCAGGAGAGAGAGCAGCT 1357
Qy      215 uLeu----- 216
Db      1358 TCTCATGACACGTGTGGCGCAGAGGGGTGTGCCCACTCAAGATGCTCTTTCATGCT 1417
Qy      217 -----GlnIleAlaArgArgLeuGlnMetTyrlGlyIleArgLeuHisProAlaLysAan 235
Db      1418 GGTGTACATCGCAGAGGCTGGAGAGATTTGGCACTGCAACATAATCCCAACACAGCAG 1477
Qy      235 gGlnGlyThrLysIleAsnLeuAlaValAlaAanThrGlyIleLeuValPheGlnGly 255
Db      1478 CTGCCAGGGCGGGGCGAGCAGAGTGGCGCAGCACTACGCAAGCTC--TTCCGTGCT 1534
Qy      255 eThrLysIleAsnAlaPheAanTrpAlaLysValArgLysLeuSerPheLysArgLys 275
Db      1535 GGTGTCAATTGGGGCC----- 1550
Qy      275 gPheLeuIleLysLeuArgProAspAlaAenSerAlaTyrlGlnAspThrLeuGlnPheLe 295
Db      1550 ----- 1550
Qy      295 uMetAlaSerArgAspPheCysLysSerPheTrpLysIleCysValGlnHisAlaPhe 315
Db      1551 -----ATCTGATCAT-CATCATTCGCG 1572
Qy      315 ePhe-----ArgLeuPheGlnGlnProLys-ProLysProLysProValLeuPhe 332
Db      1573 TTGGCTGTCTTCAACTGCTGGCAGCGCGGTGCCCAAGCTCAAGCAGCAGTGT-----T 1626
Qy      332 eArgGlySerSerPheArgPhe-----SerGlyArgThrGlnLysGlnValLeuAsp 350
Db      1627 GGCAGGCGCAGAGAGCTGCGCTGTGGAGAAAGAGCTGCCAGCAACACCAAGCTGAGAG 1686
Qy      350 yrValLysGlnGlyGlnHisLysValGlnPheGlnArgLysHisSerLysIleHis 370
Db      1687 TGGCAGCGAC-----AGCCAGTCGAGATGAGAGAGAACCCAGAGCTGAAGC 1737
Qy      370 eTrlLeaSerLeu-----AlaSerGlnProThrGlnLeuAan 383

```

```

Db      1738 GCGCGGGGCGCCCTCAAGCGCCCGGGAGCTGGGGGCGCTCATGGGGGCAAGCGGAGCC 1797
Qy      383 eRgIuValLeuGIuInSeRgInSeRThSeRLeuThRheLIuGIuValaGIuS 403
Db      1798 CCGAGGACTCGGACCTGTTCCAGAGGACACGACCTGTGAGCCGAGCGGCGAGGCC 1857
Qy      403 eRProGIuGIuIn-SeRcYeaRgRgLIuYLySgIuProlYsValSeRAlaGIuGIuPro 422
Db      1858 GAGTGGCGCCGAGGACCAAGAGAGGTGA-----CCCCA-----AACGAGCGGCC 1905
Qy      423 GlysERhIaPProSeRProAlaPProRgRgRSeRProAlaGIuYAsnLYeGIuAlaSPgLY 442
Db      1906 GGAAGCCCGACACAGCCCGCGCTTACCGGCGGCCCG-----CGCTGGCGCTCGGC 1959
Qy      443 AlaIaSeRAlaPProThGIuGIuGIuGIuValValValValSeRPaRgThRGIuInSeR 462
Db      1960 GCGGAGCTCTTCCCGCTT----- 1977
Qy      463 LysPProGIuPProProGIuPProSeRThRGIuSeRLeuThRGIuSeRProHLeuSeRGIu 482
Db      1978 -----CCCCG----- 1983
Qy      483 LeuSeRValaSeRgInGIuGIuValAlaPProAlaSeNValThRLeuSeRProaSnLeu 502
Db      1984 -----ACTTCACAGCGCGGCTTCGAGCA-----ACTCCCTCAGCTC 2019
Qy      503 SeRProaPThRlySgIuAlaSeRProLeuLeSeRProLeuLeuAsp---GIuAla 521
Db      2020 CCGCGCCGAGGGGAGGCTCAAGCCCGCTTGCGCCGCTTCCCGCCCTGAACCCCG 2079
Qy      522 CySPaRgThR-----AaPaRgLIuAspGIuGIuYARg 533
Db      2080 GCCCGCGGGGCGGGGCGGCGCTTCTGCGCCCGGAGCTCAATTAAACCGCGCGAGA 2139
Qy      534 LysaRgPhePProThaRAspLYs 540
Db      2140 CCACGCGGGGCCACAGCAAAAA 2160

RESULT 10
US-10-529-655-20
; Sequence 20, Application US/10529655
; GENERAL INFORMATION:
; APPLICANT: Ludwig Institute for Cancer Research, et al.
; TITLE OR INVENTION: Human Sarcoma-Associated Antigens
; FILE REFERENCE: L0461.70156M00
; CURRENT APPLICATION NUMBER: US/10/529,655
; NUMBER OF SEQ ID NOS: 167
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 20
; LENGTH: 2676
; TYPE: DNA
; ORGANISM: homo sapiens
US-10-529-655-20

Alignment Scores:
Pred. No.: 5.17 Length: 2676
Score: 130.00 Matches: 190
Percent Similarity: 33.4% Conservative: 118
Best Local Similarity: 20.6% Mismatches: 374
Query Match: 2.4% Indels: 244
DB: 6 Gaps: 39

US-09-555-342b-2 (1-1045) x US-10-529-655-20 (1-2676)
Qy      137 AlaLeuGIuValLYeGIuAsPLeuAlaGIuGIuYARgLeuThRcYAsnAsPThSeRAla 156
Db      48 GCGCTCCAAAGTTGCGAGAAAGCTG---CAGGCGCGCTGCGCGGAAACCGGACCTTAAG 104
Qy      157 AlaLeuLeu-----IleSeRHisIleVal 164
Db      105 AAGCATTAATGAAATATTGAGAAACTTCACCTGCTATTACAGTAGACATCTTTCG 164

```

```

Qy      165 GlnSeRGIuIleGIuAsPheAsPgluAlaLeuSeRPaRgLIuHleuAlaLYeAsnLYs 184
Db      165 GAGACTGGGCTTGGGAAAAAGTAATAGCTTG---CGAAAACAGACATGTTGAAAGC 221
Qy      185 TyrIlePProGIuInAsPAlaLeuGIuAsPlySleValGIuPHeNHisIaSnHisIle 204
Db      222 TTGGCCAGGAGCATAGTGGCCAGAGAAAGCTGTTCTCTGTGAAAGAAAT----- 275
Qy      205 GLyGIuThRProAlaGIuSeRPaRgPheGIuLeuLeuGIuIleAlaYARg----- 221
Db      276 --GCTAGCTGATGAACAGACTTTGAGAAAGCAATTCGCCAAACCGCTCGGAT 332
Qy      222 -----LeuGIuMe-TyrGIuYleARgLeuHisPProAlaLYeAsPArGI 236
Db      333 GCCCTGCAAGAGAGAGAGATGAGGGGAGACTTCACAAAGAACTTGAAAGCCAGGG 392
Qy      226 uGIuThRlySleAsnLeuAlaValAlaAsnThRGIuYleuValPheGIuGIuPheTh 256
Db      393 AGCGATCTTAAGCCCTGACCAAGCAGCAAGAAAGAAACATAGG-AAACTCGAGCTCGA 451
Qy      256 rLYeIleAsnAlaPheAsnThRAlaLYeValAlaRgLYeLeuSeRPhelysAlaRgLYaRgH 276
Db      452 GAGACTCACAAGGTCTCAAGCTCATGAGAGAGAGATGAGAGAAAGAGGTGTAC-- 509
Qy      276 eleuIleLYeAsnRgPProeRAlaAsnSeRAlaTYrGIuAsPThRLeuGIuPheLeuMe 296
Db      510 -----AGAAATGACACCACTTACTCTTCAGACCCCTGAGCTTGTATGCGCAAT 562
Qy      296 ValSeRPaRgAsPheCYeLYeSeRPhetRlySleCYeValGIuHisIaAlaPheP 316
Db      563 TCAATCCCTCCATCTTGTACAGCTCATGAGATGATGATGACACTACAGATCCCT 622
Qy      316 eARgLeuPheLIuGIuPro-----LYeProlYsProlYsProlYleuPheSeRPaR 333
Db      623 GAGAGAGACAGAGAGCCCATGTTTTCACACAGAAAGCT-----GGGAA 667
Qy      333 gGIYSeRSeRPheaRgPheSeR-----GIYARgThRGIuLYeGIuValLYeAsPty 350
Db      668 AGGCCACAGCAATGCTTTTCAGACAGACTCGGGGCGCCAGCCAAAGAACACACATGGGT-- 725
Qy      350 rValIeSgIuGIuLYeIlySlyValYLyGIuPheGIuARgLYeHisIeRlySleHisIe 370
Db      726 ---GAACCCATGGAAGAGGGTTGTGATCAAAACAAAGAGCACAATCTTCCACAA 781
Qy      370 rIleARgSeRLeuAlaSeRGIuPProThRGIuLeuAsnSeRGIuValLeuGIuInSeRGI 390
Db      782 GGAACAAGCGCCCGTGAATGCCAAGAGATGAGAAAGGCTCTGTGGAGACAGAGAA 841
Qy      390 nGIuSeRThSeRLeuThRPhelGIuGIuValaGIuSeRProGIuGIuInSeRcYAR 410
Db      842 ATCAACAAGGCTCTTCACAAAGAGAGAACCGAAGGCCACCTTCAGAGGGA-CAATGCA 900
Qy      410 g-----ArgGIuLYeSgIuPro-- 415
Db      901 GGGAGAAACCGCCCTAGTGCGGTAAAGAAAGAGAGAGACAGAGAGGGGAGAGCTCGA 960
Qy      416 ---LYeValSeRAlaGIuGIuPProGIuYSeRHisPProSeRPro----- 428
Db      961 AGAAGAAAGTGTCTCTCCACAGAGCGGCTTCAGACAAACCACTGAAAGAAAGCCAAAGC 1020
Qy      429 ---AlaPProRgRgRSeRProAlaGIuYAsnLYeGIuAlaAsPglYAlaIaSeRAlaP 447
Db      1021 ACAGAGACCCAGAAAGAACCAATTTGACAAAGACCAAGCAAGGTCTGG-ACAGCTTTGAC 1079
Qy      447 oThRGIuGIuGIuGIuVal-Val-----L 456
Db      1080 ACAGAGAAAGAGAGAGAGAGACTTGTGCCAAGTTAAAGAGAAAGGTTCTTAACAACCTA 1139
Qy      456 yAsPArRgThRGIuInSeR-LyePProGIuPProProGIuPProSeRThRGIuSeRLeuThR 475
Db      1140 AAGACTCCAGAAAGGAAAGTCAAAATTAATTGATGAACAACTACGTGGGCTCCCTC-- 1196
Qy      476 GlysERProHisLeuSeRGIuLeuSeRValaSnSeRGIuGIuValAlaPProAlaSn 495

```

```

Db      1197 -----CTTAAAGTTGAGAGACAGATATGAGAGATGAA-----TTGAGACAGCCAAACC 1244
Qy      496 ValThrLeuSerProAlaMetLeuSerProAsp-----ThrLygGlnAlaSer 510
Db      1245 ATGCTTTTGAATCTTACCTGATGATGACCAAGCCCGGAGAAAGAAAAGAAATGTTGTG 1304
Qy      511 ProLeuLeuSerProLeuLeuAspGlnAlaCysProArgThrAspAspGluAspGlu 530
Db      1305 AAAAATTACAGCCAGGACCTTGAGATTAAGACTTAAAAAATGACTCTTAAAGCACT 1364
Qy      531 GLyArgArgLybArgPheProThrAspLybAlaTyRPhelAlaLygGlnAlaSerThr 550
Db      1365 GGTAAAAAAC-----TTGACTCTCAAGTTCAAGAAATTAACCAAGTGAACAAAACC 1412
Qy      551 ----ThrGlnArgThrTyRLeuLybAspLeuGlnValIleThrSerTrpPheGlnSerThr 569
Db      1413 AAGTCAGAGAAAGCCGGCTGAGCTGATTTA----- 1442
Qy      570 ValSerLybGlnAlaAspAlaMetProGlnAlaLeuLybSerLeu----- 583
Db      1443 ----GCCAAGCTGAGAAAGAGTCCCTGATGTGTGCCAGTGTTGCCAGACCTCCCTTACC 1499
Qy      584 ----- 1lePheProAlaMetGluPro 590
Db      1500 GCGATACAGGCCAATTACCGTCCACTGCGCTTCCCTGAGCTGATATCTCTCTCCAGCCA 1559
Qy      591 LeuHisLybSerPheHisThrAsnPheLeuLybGlnIleGlnGlnAlaLeuTrpGlu 610
Db      1560 AAGCGAAAGAGGTTCTCTTCA-----CCCGAGAAAGAAAGAGTGTATTAAGTGGG 1613
Qy      611 GLyArgSerAsnAlaGlnIleArgAspTyRglnArgIleGlyAspValMetLeuLybAsn 630
Db      1614 CGCGAATGAAATTCAGATGCAGCTGAT----- 1643
Qy      631 IleGlnIlyMetLyb-----HisLeuAlaAlaHisLeuTrpLybHisSerGlnAlaLeu 648
Db      1644 ----TCTGTTTCCAAAGTGTGCTATCTCCCTAAATGATGACCTTGACCAAGCAATGATC 1700
Qy      649 GlnAlaLeuGlnAlaGlnIlyLeLybSerSerArgArgLeuGlnAlaMetPheCysArgAspPhe 666
Db      1701 CGAGTACTTAAAAACAATCATGATTCATC-----TTT 1733
Qy      669 GlnLeuGlnLybValCysTyRLeuProLeuAlaMetThrPheLeuLeuArg----- 684
Db      1734 GAAGTGGAGAGAGTCCCATCTCTGTTCTTGAACCCGTTTGGAGAGGTGACACTGAT 1793
Qy      685 ProLeuHisArgLeuMetHisTyRLeuGlnValLeu-----GluArgLeuCys 700
Db      1794 CAGCTGATCGCATAGAGAAATCAATCATGATTAATTGAAGAAACAGATCATATTATGG 1853
Qy      701 LybHisHisProProSerHisAlaAspPheArgAspCysArgAlaAlaLeuAlaGlnIle 720
Db      1854 AAAGTTTAT-----TCTCACCGAGACTTTAAGAGAAAGAACCCGGAAGATATGACTCG 1907
Qy      721 ----ThrGlnMetValAlaGlnLeuHisGlyThrMetIleLybMetGlnAlaMetPheGlnLyb 739
Db      1908 TGGGGAAGATGATGATCTGCGCTTCAAGACGCC-----CGAGACACGCG 1952
Qy      740 LeuHisGlnLeuLybLybAspLeuIleGlyIleAspAsnLeuValProGlyArgGlu 759
Db      1953 CTACGACTACTAACAAAGAAAT----- 1973
Qy      760 PheIleArgLeuGlnLybSerLeuSerLybLeuSerIlyLygGlnLeuGlnAlaArgMetPhe 779
Db      1974 ----ATCCAGTTCCGACATGCCCAATAG-----CCCAAAGGCCGACAGCAAGCAAGATGACC 2024
Qy      780 PheLeuPheAsnAspValIleLeuTyRThrSerArgGlyLeuThrAlaSerAsnGlnPhe 799
Db      2025 TTTCCTC---AAGCTGTGTGGCAAGCACTGTGACGTCCGAGAGAGGACGAAAGATTT 2081
Qy      800 LybValHisGlyGlnLeuProLeuTyRglnMetThrIleGlnGlnSerGlnAspGluTrp 819

```

```

Db      2082 GGAACGGAGAGACAGCTGTCTCTGAGAAATCAAGATCAAGCCAGCCCGGATGCCCATG 2141
Qy      820 GLyValProHisCysLeuThrLeuArgGlyGlnArgGlnSerIleIleValAlaIleSer 839
Db      2142 GGAAGACCAT----- 859
Qy      840 SerArgSerGlnMetGlnLybTyRValGluAspIleGlnMetAlaIleAspLeuAlaGlu 859
Db      2166 AGC----- 2168
Qy      860 LybSerSerSerProAlaProGlnPheLeuAlaSerSerProProAsnLybSerPro 879
Db      2169 ATCAGCTTAAATCCCAAGCCCTGAGAGCCGCTTATGATGCCCAAGACAGCAAGTGGC 2228
Qy      880 AspGlnAlaThrAlaIleAspGlnGlnSerGlnAspLeuSerAlaSerArgThrSer 899
Db      2229 CACTGGACACCAAGTGTGACAGACACTGTTCTTATGAT-----CTTAGAAACCAACT 2282
Qy      900 LeuGlnArgGlnAlaProHisArgGlyAsnThrMetValHisValCysTrpHisArgAsn 919
Db      2283 GTGAAGAAATTCGCCCAATGATGGCCAAAGCAATTAA----- 2321
Qy      920 ThrSerValSerMetValAspPheSerIleAlaValGlnGlnLeuSer----- 936
Db      2322 -----GCTTTCAGAAACAGATTCTCCCGACGATAA 2351
Qy      937 -----GlyAsnLeuLeuArg-----LybPheLybAsnSerArgGly 948
Db      2352 ACTGAGACTTGCTGCTTGGAAATGAAATCTGGAGGCGAGAAATCAAGCAAGTGGGGGT 2411
Qy      949 Trp 949
Db      2412 TGG 2414

RESULT 11
US-10-576-496-45
; Sequence 45, Application US/10576496
; GENERAL INFORMATION:
; APPLICANT: Teijin Pharma Limited
; APPLICANT: Komori, Tohshisa
; APPLICANT: Kanacani, Naoko
; APPLICANT: Yoshida, Carolina Andrea
; APPLICANT: Zanna, Akira
; APPLICANT: Kobayashi, Shinji
; APPLICANT: Yamana, Kei
; TITLE OR INVENTION: BONE AND/OR JOINT-DISEASE-ASSOCIATES GENES
; FILE REFERENCE: Q94468
; CURRENT APPLICATION NUMBER: US/10/576,496
; CURRENT FILING DATE: 2006-04-20
; PRIOR APPLICATION NUMBER: PCT/JP2004/015879
; PRIOR FILING DATE: 2004-10-20
; PRIOR APPLICATION NUMBER: JP 2003-359172
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 45
; LENGTH: 3543
; TYPE: DNA
; ORGANISM: human
; NAME/KEY: CDS
; LOCATION: (373) .. (3003)
US-10-576-496-45

Alignment Scores:
Pred. No.: 11
Score: 127.00
Percent Similarity: 33.8%
Best Local Similarity: 19.4%
Query Match: 2.3%
DB: 6
Length: 3543
Matches: 138
Conservative: 102
Mismatch: 278
Indels: 193
Gaps: 31
US-09-555-342B-2 (1-1045) x US-10-576-496-45 (1-3543)

```


385 ValLeuGlnGlnSerGlnSerThrSerLeuThrPheGlyLys----- 399
 18 CTGGAGAGCGGGCTCCAGAGAGGCTTCTGGGCGCCGAGAGAGGAGCGGGTCCCAAGAGG 77
 400 -----GlyAlaGlnSerProGlyGlyLysSerCysArgArgGlyLysGlu---ProLys 416
 78 GTGGCTCTGTGGAGAGCTCCGGGGGTGGGGGGCGGGCGCTCCGGGAAGAAGCTTCCGAG 137
 417 ValSerAlaGlyLysProGlySerHisPro----- 426
 138 GTCCCCCGCCCGTCACGTGGGGCGCGCGCCCGCCGCTGCGGTCCGCTGGTGGTTC 197
 427 -----SerProAlaProArgArgSerProAlaGlyLys 437
 198 GGGGCGCTGGTCCGCGACGTTGGTCCGTGGGCGCAGTGGCGCGTGGCTCTTGGGGCTCT 257
 438 Lys-----GlnAlaAspGlyValAlaLaserAlaPro-ThrGlnGlnGlnGlu---GluVal 454
 258 CATGTTTGAAGGTGGAGAGGACACCGAGACCGCGCCGACACCTGAGCCGCCCGAGAGGA 317
 454 ValLysAspArgThrGlnGlnSerLysAspProGlnProGlnProSerThrGlySerLe 474
 318 GCCTCGCGCCCGTACCCAGTAAAGAGAGAGGAGCGCCAGCGCAAAAGAGTATGAGC 377
 474 uThrGlySerProHisLeuSerGlnLeuSerValLaserGlnGlyValAlaPro-- 493
 378 TTCTGATGCTAGTACGTGGCGCTGGAGAGCTGCCGAGCAAAATGAGCGGAGATCATGGCAG 437
 494 -----AlaAsnValThrLeuSerProAsnLeuSerProAspThr 506
 438 CACTAAACAGGTGACAGATCTTAGTGAATGTACTGTGAGCTGAGCTGGCTTCCCGCGC 497
 506 rLysGlnAlaSerProLeu---LaserProLeuAsnArgGlnAlaCysProArgThr 525
 498 CTCTACATGAACCCCTTCCGCGTCTCCATCTCATCGAGAGCTTGAGGCTGGCTTGG-- 555
 525 rAspAspGlnuArgLysArgArgLysArgPheProThrAspLysAlaTyrrPheAla 545
 556 -----GAGATGCTGAGCTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 596
 545 AluGlnAlaSerThrThrGlnArgThrTyrrLeuLysAspLeuGlnValLleThrSerTr 565
 597 C---CAGATCCCTGGGCCAAGACAGCTGCCTACATAAAGAA-----TG 635
 565 rPheGlnSerThrValSerLysGlnAspAlaMetProGlnLysLeuLysSerLeuLlePh 585
 636 GTTGAAGAGAGCTGTCCAGGTAAACAC----- 666
 585 eProAsnPheGlnPheLeuHisLysPheHisThrAsnPheLeuLysGlnLleGlnGln 605
 667 -----CAGAGTGTCTAGTAAAT-----GAAACCTTACAGAGAGAG 701
 605 gLeuAlaLeuTrpGlnGlyArgSerAsnAlaGlnLleArgAspTyrrGlnArgLleGlyAs 625
 702 CTTCGACAGTGAAGAGGAGTAAAGAGTCCCTCATATTGACAGTGAATGCTGCACAGA 761
 625 rValMetLeuLysAsnLleGlnGlyMetLys-----HisLeuAla 639
 762 CCAAGTA-----GAAAGCCAGAGGAGAAAGATTGAGACCTGGAAGTGTCTGGAAGG 815
 639 HisLeuTrpLysHisSerGlnAlaLeuGlnAlaLeuGlnLysGlnLleLysSerSerArg 659
 816 ACACCAAGGTAAATCTCAATGCTGCTGAAGAGATGCTTCAACAGAGAGCTGTAAGCCGAC 875
 659 gArgLeuGlnuAsnPheCysArgAspPheGlnLeuGlnLysValCysTyrrLeuProLeuAs 679
 876 ATCTCTTAAGACCCAGAGAGCTGATCTG---ATGACTAAAGTGTGTAGCTGAAGCTCAA 932
 679 nThrPheLeuLysArgPro----- 685
 933 GCTGGTGGCATGAG 992

686 ----LeuHisArgLeuMetHisTyrrLysGlnValLeuGlnuArgLeuCysLysHisLlePr 704
 993 GTTACTGCAGAGAGCTCAGACACCTTAATTAAGTGAAGAGTGGAAATGAAGAGAA 1052
 704 oProSerHisAlaAspPheArgAspCysArgAlaAlaLeuAlaGlnLleThrGlnMetVal 724
 1053 TCAGTAAATGAAGAGTAAAGGCCATGAAGCTGAATGCCCGCAGCTGCAAAACAGAGT 1112
 724 Ala-----GlnLeuHisGlyThrMetLys----- 733
 1113 GGGCTGAAGATGACAGAAATTAAGCTGTGCAGACCGAGCTCTCCGAGACAGAGCTCT 1172
 734 -----MetGluAsnPheGlnLysLeuHisGlnLeuLysLysAspLeuL 748
 1173 CCACATGAGATGACACACAGAGAGAGACCAAGAAATTCAGAGCTGGA-----AT 1223
 748 eGlyLleAspAsnLeuValValProGlyArgGlnPheLleArgLeuGlySerLeuSerLys 768
 1224 GGGATGGAACCTTGTGCTTGGCCATGAAGATTAAGAGACGCT-----CGGATGAGGA 1277
 768 eLeuSerGlyLysGlyLeuGlnGlnArgMetPhePheLeuPheAsnAspValLeuLeuTy 788
 1278 GCTTACGGGGCTGTTAAACAGTACCGGAG-----GTAAAGAGATGTGATGCT 1328
 788 rThr-----SerArgGlyLeuThrAlaSerAsnGlnPheLysValHisGlyGlnLe 805
 1329 CACTCAGAGGCTTCCGAGAGAACTCTCATATCATATAAGAAACCGAGAGAGCTTT 1388
 805 uProLeuTyrrGlyMetThrLleGlnGlnSerGlnuAsp-----GlnTyrrGlyValPr 822
 1389 CAGAAAGTGAACGCTCAAAATTAAGACCTGAAAGATTAATTAACAAGAGATGCTTCC 1448
 822 HisCys-----LeuThrLeuArg 828
 1449 AAGATGTAGCTCTCTTCAAGTGGGCGCCCTCCATTCGACAGAAATCATCGAAACAG 1508
 828 gGlyGlnArgLysSerLleLleValAlaLaserSerArgSerGlnMetGlnTyrrPva 848
 1509 GGCCTCAAGAAAGCTCTTGTAGTCTTGAAGACTTGAAGAGTGAATGTGTGGATGAAGT 1568
 848 LgluAspLleGlnMetAlaLleAspLeuAlaGlnLysSerSerSerProAlaProGlnuPh 868
 1569 TATGATGAGAAACAGAGCCCTCCCGGTGTGAACCCAGAGAGAGCCCT-----TT 1619
 868 eLeu----- 869
 1620 CTGGCGGAGACAAATATCCACTTATCTGGGAAGCTTTCAGAGGCCAGGCCCAATGG 1679
 870 ----AlaserSerProProAspAsnLysSerProAspGlnAlaThrAlaAla----- 885
 1680 AGAGGCTGCAAATCTCTCCACCATGTGCCAGCTTACCCAGGGAGACAGCTGCT 1739
 886 -----AspGlnGlnSer-----GluAspAspLeuSerAl 895
 1740 GAGGCTGAGAGACACAGAAAGTGGCTGGAGAGACAGACTGTGTGCATGACCTTCAATC 1799
 895 aSerArgThrSerLeuGlnArg-----GlnAlaPro-----HisArgL 908
 1800 CACATCAACGGGCACTGAATCAGGCTCTCATGCTCTCTGACACCAAGATGTAACGGAA 1859
 908 yAsnThrMetValHisValCysTyrrHisArg----- 918
 1860 TCCCAAGAGCTTAAGAAATCTTGGGAAATATCCGAAGAACTCAGTACGAAATTTCTA 1919
 919 -AsnThrSerValSerMetValAspPheSerLleAlaValGlnuGlnLeuSerGlyLys 938
 1920 CACTGACACAGCTGGAGATGGCAGAGTTTCAGAGAGTGGGCTCCGGGCAACCGGAGGCC 1979
 938 nLeuLeuArgLysPheLysAsnSerAsnGly 948
 1980 AAGACTCTTAGAGACAGAGAGCTCAAGGA 2010

RESULT 12

```
US-60-796-903-4
; Sequence 4, Application US/60796903
; GENERAL INFORMATION:
; APPLICANT: Johnston, Patrick
; APPLICANT: Harkin, Paul
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Colorectal Cancer Microarrays and Methods of Use
; FILE REFERENCE: 55815-300 (327365)
; CURRENT APPLICATION NUMBER: US/60/796,903
; CURRENT FILING DATE: 2006-05-02
; NUMBER OF SEQ ID NOS: 1071
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 4
; LENGTH: 6373
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-60-796-903-4

Alignment Scores:
Pred. No.:      23.8      Length:      6373
Score:          126.50    Matches:      144
Percent Similarity: 37.2%  Conservatve: 104
Best Local Similarity: 21.6% Mismatches:    243
Query Match:     2.3%     Indels:       177
DB:              8        Gaps:         32

US-09-555-342B-2 (1-1045) x US-60-796-903-4 (1-6373)
QY      343   ThrGlnLysGlnValLeu---AspTyrValLysGlnGlyGly----- 355
DB      897   ACACCTAAGCAGATCTCTCAATTAACCTCAAGAAAGAACACTGATATGACGATGCT 956
QY      356   ---HisLysValGln-----PheGluArgLysHisSerLysIleHisSerLysArg 372
DB      957   ACTAAAGAAAGTCAACTTGTCTTACG-----AAATCCAGACCTTAAAG 1004
QY      373   SerLeuAla-----SerGlnProThrGlnLeuAsnSerGlnVal 385
DB      1005  TCTGCAGCAGCTGGAGCGCCCAAGAAATACCAAGCTTAAATTCACATCTGAAGTC 1064
QY      386   LeuGlnLysGlnGlnSerThrSerLeuThrPheGlnGlyGly-----Ala 401
DB      1065  AAAATCTATTGTAACAGAAAGAAAGTTCACATAGA-AGTGCCTGAACCTTCAGAGACA 1123
QY      402   GluSerProGlnGlnSerCys-----ArgArgLysLysGlnProLysVal 417
DB      1124  GCTTCAACTGGAAGTCAAGATCAACAGACATTTAAAGAGAGAGAGAGATGAG--- 1180
QY      418   SerAlaGlnGlnProGlnSerHisProSerProAlaProArgSerProAlaGlyAsn 437
DB      1181  -----AGCAAACTTGAAGAGCTCCG----- 1201
QY      438   LysGlnAlaAspGlyAlaAlaSerAlaProThrGlu-----GlnGlnGlnGluVal 454
DB      1202  -----AAGCCAAACACAAAGGAGGAGGAACTCCAC 1237
QY      454   IValLysAspArgThrGlnGlnSerLysProGlnProProProSerThrGlnSerLe 474
DB      1238  ATTCCAGCAACGACTCGAAGAAAGTGAAGGGGAGCT-CCGAAAGAAATCTGGAGAGCTAT 1296
QY      474   uThrGlnSerProHisLeuSerGlnLeuSerValAsnSerGlnGlnGlnValAlaProAl 494
DB      1297  TCCAGTGAAGATGGAACGGAGACAGCATCAGACTGAGATCAGAGATCTCCAGAGCCAGC 1356
QY      494   e-AsnVal-----ThrLeuSerProAsnLeuSerProArgThrLysGlnAlaSerP 511
DB      1357  TCTCAAGAAATGACAGATGAACCTGACAGTGCAGAAACGATTCGAGAGACAGGAGAGAG 1416
QY      511   rLeuLysSerProLeu-----LeuAsnAspGlnAlaCysProArgT 525
DB      1417  CTCGATTTGAGGAGCTCTTACAGGCAAAACAGATCTTCAAGATCTGCTGATTCGCAAG 1476
QY      525   hrAspAspGlnAspGlnGlyArg---ArgLysArgPheProThrAspLysAlaTyrPheI 544
```

```
DB      1477  AGGAGCAAGAAAGACCTTGTGAAGAAAGCGAGCGTGAATCACCGCCCTGAAGGAGGCC 1536
QY      544   LeAlaLysGlnValSerThrThrGlnArgThrTyr----- 555
DB      1537  TGAAGAAAGAGTTTCCAGCATGATGAGATGGAACAAGCTGAAGAGCAATATGATG 1596
QY      556   -----LeuLysAspLeuGlnValIleT 563
DB      1597  CTGAGTTCAAGCCCTTGAAGAGATGGAAGAAAGCAACCAAGAAATGTGAGGCTTTGG 1656
QY      563   hrSerTrpPheGlnSerThrValSerLysGlnLysPhe---AlaMetProGlnAlaLeuLys 582
DB      1657  CGAGC-----AGAGCAACACTTCAGAGCAAGACACAGCGCGGAGCTGAATGCGCGCTGA 1710
QY      582   eLysLysPheProAsnPheGlnProLysLysPheHisThrAsnPheLeuLysGlnI 602
DB      1711  AGCTTTCAGAGAGAGATGAGAACTGCAG------GGAAGAAAGCAAGAGC 1758
QY      602   LeGlnGlnArgLeuAlaLeuTrpGlnLysArgSerAsnAlaGlnIleArgAspTyrGlnA 622
DB      1759  TGAAGCGGAGAGTTGCT-----CAGCTTCAAGGCAAGATGAGAGACCTGAAA- 1805
QY      622   rGlnLysLysPhe-----ValMetLeuLysAsnIleGlnGly----- 633
DB      1806  ---GGCGATGAAGCCCAAGCGGAGGAAAGCGCTGAAGAAAGTACGAGGAGAAATACGAC 1860
QY      634   --MetLysHisLeuAlaAlaHisLeuTrpLysHisSerGln-----AlaLeuG 649
DB      1861  AGTTAGAGAGAGCCCTTGTGTGACGCCAGAAAGAGAAAGAAAGAAAGCTGTGTACGCCAGAA 1920
QY      649   IuAlaLeuGlnAsnGlnLysSerSer----- 658
DB      1921  GGGCCCTGAGAAATGAATCTGAGAGCTGTCAAGGGAATCTGAGTCAAGTACCAAGAGC 1980
QY      659   --ArgArgLeuGlnAsnPheCysArgAspPheGlnLeuGlnLysValCysTyrLeuProL 678
DB      1981  AGAAGCAGTGTGTGAGAACTCAAGAGAGAGAGTGAAGAG------ 2024
QY      678   euAsnThrPheLeuLeuArgProLeuHisArgLeu-----M 690
DB      2025  -----GACAGCTTAAGAAAGTGAAGAAACAGATGAGAGATGAGCGCT 2067
QY      690   eHisTyrLysGlnValLeuGlnArgLysCysLysHisHisProProSerHisAlaAsp 710
DB      2068  GGACCTCGGGCAAAACCATGAGAACTGCAGAAAG-----GAGATGCACAGCA 2115
QY      710   heArgAspCysArgAlaAlaLeuAlaGlnLysLeuThrGlnMetValAlaGlnLeuHisG 730
DB      2116  TT-----GTTGAGGCGCTCCCGTAACTCAACCTGTGAGCTCCAGAACACAGCTG- -G 2163
QY      730   hrMetLysLysMetGlnAsnPheGlnLysLeuHisGlnLysLysValAspLeuLysLeG 750
DB      2164  ATGAGTTTAAGAGAAACCGCAGGAGCTCGCAGAAATGCAAAAGCAGCTTG- - 2216
QY      750   LeAspAsnLeuValValProGlnArgGlnPheLeuArgLysSerLeuSerLysLeuS 770
DB      2217  -----AAGGAGAAACGCTGAGAGGCAAGAAAGTCCGAGCTGA 2253
QY      770   eGlnLysGlnLysGlnGlnArgMetPhePheLeuPheAsnAspValLeuLeuTyrHis 790
DB      2254  CAGCGATGAAGAAATGACAGATGAGATGCGTGTGATGAGAAAGTTTCGGAGACTACAGCA 2313
QY      790   eArgGlnLysLeuThrAlaSerArgAsnGlnPheLysValHisGlnLysLeuProLeuTyrG 810
DB      2314  GAGCT-----CAGGATGAGACCTCACAAAGACAGCTTCTGAGAC 2355
QY      810   eThrLysGlnGlnSerGlnAspGlnLysTrpGlnValProHisCysLeuThrLeuArgLys 830
DB      2356  AGAGCTGAAGAGCTGAGATATGAGCTGGAAGCAAGAGT-----CAGCTCAAGAGATG 2409
QY      830   IuArgGlnSerLysLysValAlaAlaSerSerArgSerGlnMetGlnLysTrpValGlnA 850
```

```
Db 2410 ACCGACAGAGCTGCTC-----AGCAGATGAGGACCAAGTGTCTC 2451
Qy 850 epIleGImeTAlAllepleuAlaGluYSeSeSeSeProAlaProGluPheLeuA 870
Db 2452 AACTGGAGATGAACTGGAGAAAGAGAAACAACTCAGATTTCGTCTGTAGAGAACTCA 2511
Qy 870 lAsSeSeProProApeAnlySeSeSeProApeGluAlaThrAlaAlaApeGIngluSeRg 890
Db 2512 GTAGAGACGAGGAAACAGATGAGCAGGTGAGAAATGACTTCTCAGAGAGAGCTGCCA 2571
Qy 890 lASeApePLeuSeSeAlaSeSeRgThrSeSeGluAlaGlnAlaProhIsaRgIyAeNT 910
Db 2572 GACAAGACTTGCAGAGTGCACAAAGATTTCCTGTGAGAGCAGAAACAAGACTTAAAGAGCC 2631
Qy 910 hmeVAlhIeVal 914
Db 2632 GGATTATCCACCTG 2645

RESULT 13
US-11-366-965-1/c
Sequence 1, Application US/11366965
GENERAL INFORMATION:
APPLICANT: Griffiths, Remy
APPLICANT: Holseth, Susan K.
APPLICANT: Zagursky, Robert John
APPLICANT: Metcalf, Benjamin J.
APPLICANT: Peek, Joel A.
APPLICANT: Sankaran, Banumathi
APPLICANT: Fletcher, Leah Diane
TITLE OF INVENTION: CHLAMYDIA TRACHOMATIS POLYNUCLEOTIDES AND VECTORS, RECOMBINANT HC
FILE REFERENCE: GEN-T109X
CURRENT APPLICATION NUMBER: US/11/366,965
CURRENT FILING DATE: 2006-03-02
PRIOR APPLICATION NUMBER: US/09/201,228
PRIOR FILING DATE: 1998-11-30
PRIOR APPLICATION NUMBER: US 60/107,077
PRIOR FILING DATE: 1998-11-04
PRIOR APPLICATION NUMBER: FR 97-16034
PRIOR FILING DATE: 1997-12-17
PRIOR APPLICATION NUMBER: FR 97-15041
PRIOR FILING DATE: 1997-11-28
NUMBER OF SEQ ID NOS: 5982
SOFTWARE: PasteSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 1038608
TYPE: DNA
ORGANISM: Chlamydia trachomatis
FEATURES:
NAME/KEY: misc_feature
LOCATION: (1)...(1038608)
OTHER INFORMATION: n = A,T,C or G
US-11-366-965-1

Alignment Scores:
Pred. No.: 1,11e+04 Length: 1038608
Score: 126.00 Matches: 157
Percent Similarity: 35.1% Conservative: 108
Best Local Similarity: 20.8% Mismatches: 269
Query Match: 2.3% Indels: 221
DB: 7 Gaps: 38

US-09-555-342b-2 (1-1045) x US-11-366-965-1 (1-1038608)
Qy 324 ProlyApePro-----LysProValLeuPheSeSeRgIy 334
Db 752301 CCGAGCCGGGTTCTGTGATGAAACAGTATCACTGCCCTGTTTGTATAAAGATGGG 752242
Qy 335 SerSeSePhe-----ArgPheSeSeGlyAArgThhGlnlySeGln 346
Db 752241 GTTTCCTTTTGAATGGAAGCTGCCAAGCTCTGTTTTTAAAGGGGAAAGAGTAGAAGCA 752182
Qy 347 -----ValLeuApeTyValIySeGluGluGlyhISeIys----- 357
```

```
Db 752181 GATATAGAGTTATGCTTCACAGACAGAGTGGTATATGAGAGAGCTTCAGCGCAAGGC 752122
Qy 358 -----LysValGlnPheGlu-----ArgIyhiSeSeIySllleISeRtleArgSeR 373
Db 752121 TCTTCACGCGTTTCCTCCGAAAGTCTCAAAAAGAGAGCAAGAAATCCAAAGTTCCTTCG 752062
Qy 374 LeuAlaSeRgInProThhGlnLeuApeSeSeGluValLeuGlnGlnSeRgInSeRgInSeR 393
Db 752061 GAGCAAAAGTTTCTGAGAGAGTATCTAGATTCAGCAATAGCGAAAGAAACAGATTAGCG 752002
Qy 394 SerLeuThrPheGlyGluGluValAlaGluSeSeProGlyGlnISeSeCyAArgIyLyS 413
Db 752001 GTGTTTTTTTACGTTCTGTTGAGAGAGAGCGCTGA-----ACACCAAAAGAACTCTCT 751948
Qy 414 GluProlyValSeRAlaGlyGluProGlySeRhiSPSeSe----- 427
Db 751947 GAGCCAAAGTCTCTTCAAGAA-----GACACAGACTCTTCTGTTCAGAGAAAAAAG 751891
Qy 428 -----ProAlaProAArgSeSeProAlaGlyApehlySeGlnAlaApeGlyAlaAlaSeR 445
Db 751890 GATCTTGAGCTTCTTGGCAAGTCAAGACAACTTAAACAACTCACTCCATCCAGCGCAGT 751831
Qy 446 Ala---ProThhGlnGluGluGlu---GluValIyAlaApeAArgThhGlnGlnSeRlyS 463
Db 751830 GGTGAACCAACCAATCTCAAAACGCGAGTATGAGAAAGAAACAGAAACATCCGCCGATCA 751771
Qy 464 ProGlnProProGlnProSeRThhGlySeRleuThhGly-----SeRProhISeSe 481
Db 751770 AATCAGCGCCACAGCTTCTTCTGCTTCACAAAATCGGCTTCAAAAGTCCGAAAAATCG 751711
Qy 482 GluLeuSeSeValApeSeRgInGlyGlyValAlaProAlaApeValThleuSeR----- 499
Db 751710 GAGCAACAACCTTCAAAAGCTTCCCAATCCCGGAAACTCCAGAGCCGTCAGAGAA 751651
Qy 500 ProApeSeSeSeProApeThhGlyGlnAlaSeSeProleuIleSeSeProleuApehlyS 519
Db 751650 CTTAATAGCGGTAGAGAAAGAAATCCGCCATCTCCAAATGAGAAAGCTTCGTAACAGAA 751591
Qy 520 GlnAlaCyProAArgThhApeApeGluApe---GluGlyAArgIyAArgPheProThh 538
Db 751590 GAAAGCAGCTCAGGAGCAGAGTGAAGAAAGAGAGGT----- 751552
Qy 539 AspIySAlaTyRPhelleAlaIySeGluValSeRThhThhGlnAArgThhTyLeuIyApe 558
Db 751551 -----GAAAGAGATCTGCTGAAAGCGCAGCAAAATGAAG 751516
Qy 559 LeuGluValIleThhSeRThhPheGlnSeRThhValSeRlyGluApeAlaMeSeProGlu 578
Db 751515 CTAAAGCAGAGCGTCT-----CAAGAAAGAGAAAGAAAGAAATGAANNAGAG 751462
Qy 579 AlaLeuIySeSeLeuIlePheProApePheGlnProleuIySeRPhenIeThhApe-- 597
Db 751461 GTTCTTGCTCCCTT-----AATGTCAAGATCTTTCGTTTGATCAAGAAATC 751411
Qy 598 PheLeuIySeGluIleGluGln----- 604
Db 751410 TTCCCTGCTGAGATGAAGATCTTGACAGAAACAGTTGCGGTGATTTGACGCAACCA 751351
Qy 605 -----ArgLeuAlaIeUtr 609
Db 751350 TCACGATTTTGTGAAGTCTTGTGTCGGAATATCGGTGCTGAATTCATTTGGAT 751291
Qy 610 GluGlyAArgSeRAlaGlnIleAArgApeTyRglnAArgIleGlyApeValMeSeIyS 629
Db 751290 AGTGGAAACCTATATCTATGTAAGATGATCCGAG-----GTTGACACATGCTTAAAGT 751234
Qy 630 AsnIle-----GlnGlyMeSeIySeIleuAlaAlaIleSeRThhIySeSe 645
Db 751233 GATATGAGTATTTGCGCCCAACATGCCAAGATCACTTATCGGCAT-----GATATTTCA 751180
Qy 646 GluAlaIeGluAlaIeU-----GluApeGly---IleIySeSeSeRArgIyGluGln 662
```

```
Dh 751179 GTTTGATGAAAGATCTGGTAGTAAAGATGCGGATGTTGTAAGGCGCAAGATTGAA 751120
Qy 663 Asn-PheCysArgPheArgPheGluLeuGlnValCys-----TyrLeuProLe 678
Db 751119 CATCAATCTACGCTCTGTCGCAATCAAGTTGCTCTAGAACAAAGTTATTTCTACTT 751060
Qy 678 uAsnThrPheLeuLeuArgProLeuHis-----ArgLeuMetHis 691
Db 751059 GTGACATTAATCTGCTCTGTCGCAATGAGTAAAGGAGATTTCTTCGAAGATTAAAGG 751000
Qy 691 eTyrLysGlnValLeuGluValGluCysValArgHisPheProSerHisAlaAspPheArg 711
Db 750999 TTAT---TTGCTCTCTCGCAATCTCTCGAAGAGATTGCTGCGCAAGCTCGCAAGAGAA 750943
Qy 711 gAspCysArgAlaAlaLeuAlaGluValLeuThrGluMetVal-----AlaGlnLeu 727
Db 750942 GAAGAGAAAGCAAAACGCTGCTACGTTGCCAACAGAGTCTTTATTAATTACCTTTCTTCA 750883
Qy 727 uHisGlyThrMetIleuMetIleuMetGluAsnPheGlnLeuLeuHisGluLeuLysAspLe 747
Db 750882 GGAGCGTTAGCTGCTGCTCTTGGAAATAGAACAACTTTGTTCCATAGC-AGGAAAGT 750824
Qy 747 uIleGlyIleAspAsnLeu---ValValProGlyArgGluPheIleArgLeuGlySerLe 766
Db 750823 AGTTCTATATGATCAAAATCGATTGATTCATGATTAATGAAATGCTTA----- 750778
Qy 766 uSerLysLeuSerGlyLysGlyLeuGlnGln-----ArgMetPhePheLeuPheAsnAs 784
Db 750777 -----ATTCAAGCACTTTCCAACTGACGTTTACGTTTCAAT-- 750742
Qy 784 pValLeuLeuLysThrSerArgGlyLeuThrAlaSerAsnGlnPheLysValHisGlyGly 804
Db 750741 -----AAGAAACATGAGCA 750728
Qy 804 nLeuProLeuTyrGlyMetThrIleGluGluSerGluAspGluTyrGlyValProHisCys 824
Db 750727 GTTGTCTTATATGGGCACTGA----- 750706
Qy 824 sLeuThrLeuArgGlyGlnArgGlnSerIleIleValAlaAlaAspSerArgSerGluMe 844
Db 750705 -----KAAATAGCAAT--GATTAAGACGAGTT 750680
Qy 844 tGluLysTyrValGluAspIleGlnMetValAlaAspLeuAlaGluLysSerSerSerPr 864
Db 750679 ACTTACAAAGTGATGCTCTCTCG----- 750655
Qy 864 oAlaProGluPheLeuAlaSerSerProProAspAsnLysSerProAspGluAlaThrAl 884
Db 750654 -----TTGTCAAGTCGGTAGATGATGATGATGATGATGATGATGATGATGATG 750608
Qy 884 aAlaAspGlnGluSerGluAspLeuSerAlaSerArgThrSerLeuGluArgGlnAl 904
Db 750607 GCAGAGATGAACCTTCTCTGTTGTTAAATCCAGAAATTTAAAGTTATCAGCATGCAATC 750548
Qy 904 aProHisArgGlyAsnThrMetVal----- 912
Db 750547 TCCAGACCGCGGATTTTGTAAATCAGCGGATATCTAAAGACAGAAAGAACAGCAGCTTG 750488
Qy 913 -----HisValCysThrPheIleArgPheThrSerValSerMetValAspPheSerI 929
Db 750487 TTGGCTGATATTAATCTAATCTAATTTTAAT---TACCTTCACTATGTTGATTAATAGG 750431
Qy 929 eAlaValGluAsnGln-----LeuSerGlyAsnLeuLeuArgLysPheLysAs 945
Db 750430 GATTATGAATCACAAGCTCATGAAGCTCTTCTGAGACATCTTTTGCA----- 750382
Qy 945 nSerAsnGlyTyrGlnLysLeuTyrValValPheThrAsn 958
Db 750381 ----TCAGGTTTTCGAAACGTTCATGTGCTCTTACCAAT 750346

RESULT 14
PCT-US06-14800-71
; Sequence 71, Application PC/TUS0614800
```

```
; GENERAL INFORMATION:
; APPLICANT: Children's Hospital Medical Center
; APPLICANT: Wong, Hector R.
; APPLICANT: Aronow, Bruce J.
; APPLICANT: Shanley, Thomas P.
; TITLE OF INVENTION: METALLOTHIONEIN AS AN EARLY BIOMARKER
; TITLE OF INVENTION: FOR DEATH SECONDARY TO SEPTIC SHOCK AND AS A NOVEL
; FILE REFERENCE: CHM027.001VPC
; CURRENT APPLICATION NUMBER: PCT/US06/14800
; CURRENT FILING DATE: 2006-04-19
; PRIOR APPLICATION NUMBER: US 60/673,656
; PRIOR FILING DATE: 2005-04-20
; NUMBER OF SEQ ID NOS: 120
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 71
; LENGTH: 6675
; TYPE: DNA
; ORGANISM: Homo sapiens
PCT-US06-14800-71

Alignment Scores:
Pred. No.: 31 Length: 6675
Score: 125.00 Matches: 162
Percent Similarity: 32.2% Conservative: 80
Best Local Similarity: 21.5% Mismatches: 279
Query Match: 2.3% Indels: 232
DB: 1 Gaps: 35

US-09-555-342B-2 (1-1045) x PCT-US06-14800-71 (1-6675)
Qy 285 AsnSerAlaTyrGlnAspThrLeuGluPhe-----LeuMetAlaSerArgAsp 300
Db 3593 AACTCAGCCAGAAACCTCTCTCCAGCTCAGCCAGCCAGCAAGCTCTCTCCAGAACTCA 3562
Qy 301 PheCysLysSerPheThrLysIleCysVal-GluHisAlaPhePheArgLeuPheG 320
Db 3653 TTCAGAGAAACCTTCCCGCAGCCCTCGTCAGATGCCAT--TTCTC 3697
Qy 320 uGluProLysPro-----LysProLysProVal-----LeuPheSerArgG 334
Db 3698 CAGACCTCAGCCATACAAACCTTCTCCAGACTCAGCCATACAAACCTTCTTTTACACC 3757
Qy 334 ySerSerPheArgPheSerGlyArgThrGlnLysGlnValLeuAspTyrValLysGlu 354
Db 3758 TCAGCCAGAAACCTCTCTCCAGAACTGCTCA-----GACAACTTCTCCAG 3808
Qy 354 yGlnHisLysLysValGlnPheGluArgLysHisSerLysIleHisSerIleArgSerLe 374
Db 3809 CCTCGGTCAGATGCCCTTCTCCAGACCTCAGCCATACAAACCTTCTCTAGA----- 3863
Qy 374 uAlaSerGlnPro-ThrGluLeuAsnSerGluValLeu-----GluG 388
Db 3864 -----CTTCAGCCAGAAACCTCTCTCCAGAACTCAGCCATTAATGACTCTCTCCAGAAC 3919
Qy 388 InseGlnGlnSerThrSerLeuThrPheGlyGluGluAlaGluSerProGlyGlyGlns 408
Db 3920 TCAGTCAGAAACCTTCTCCAGCCCTCGTCAGATGCCATTTCTCCAGACT-CAGC 3978
Qy 408 eCysArg-----ArgGlyLysGluProLysValSerAlaGlyLysGluProGlySerH 425
Db 3979 CATACAAACCTTCTCTAGACTTCAGCCAGCAAAACCTCTCCAGAACTCAGTCAACA 4038
Qy 425 iAspProSerProAlaProArgArg---SerProAlaGlyLeuLysGlnAlaAspGlyAla 444
Db 4039 AACCTTCTCCAGACCTCGGTCAGATGCTCTTCTCCAGACCCAGCCATACAAACCTT 4098
Qy 444 iAspSerAlaProThrGlnGluGluGluValValLysAspArgThrGlnGlnSerLysP 464
Db 4099 TCTCTAGACTTCAGCCAGCAAAACCT-----CTCTCAGAACTCAGTCAAGCAAAAC 4149
Qy 464 rGlnPro-----Pro-Gln 468
```

```

Db 4150 CTTTCCCGAGACCTGAGTGAATGCCCTTTTGCAGATCTCATGAAATTCCTCCCTTACC 4209
Qy 469 ProserThrglySerLeuThrglySerProHleuSerGlu-----LeuSerValAen 486
Db 4210 CCAAGCCTCGACGACATGACACTTCTCCAGACCTTGGTGGAGACAGATCTTCCCAAC 4269
Qy 487 ---SerGlnGlyValAlaPro-----AlaAenValThrlLeuSerProAenLeuSer 503
Db 4270 TTTGGTCAGATGATCCCTTCCCGACAGCTCCAGCCAGGTGATCTTCTTCCAGACATCACT 4329
Qy 504 -----ProAerThrlGlySerGlnAlaSerPro----- 511
Db 4330 GACACACCTTCTCCCGATCTCGACAGATATCATCTCCCTCCAGACCTTGATCAGATA 4389
Qy 512 -----LeuLeu----- 513
Db 4390 TTCTACCTTCTGAACTTAGTCAGTCAATGCTTCTTCAAGAAATTAATGAGCTTTTCT 4449
Qy 514 -----SerProLeuLeuAenAerP----- 519
Db 4450 TATCCAGACCTTGGTCAGATGCCATCTCTTCATCTCTCACTCAATGATATCTTTCTA 4509
Qy 519 ----- 519
Db 4510 TCAGAGAAATTAATCCACGTGTTATAGTGGGCTCAGTAAAGATGTACAGATTACAT 4569
Qy 520 GlnAlaCyProAerG-----ThraAerGluAerGluGlyAerGlyAerG 535
Db 4570 GAGATCATCTCCAGAGAGAGCTCCAGACAGTCAAGATGACTATGCTGAATTTGATTAAT 4629
Qy 536 PheProThraAerGlyValAerPheHleuAlaGlyValAerThrlGluAerThrlGly 555
Db 4630 GTGCCCTATGATGACCCCTAC-----AAACGATGTTAAGGACAAAC 4671
Qy 556 Leu-----LysAerLeuGluValAlaThrlSerThrlPhe----- 566
Db 4672 ATCAACCTCCAGAGATCTCGACAAACATTCAGCATGTGTCCTCCGACAGACAAATGGA 4731
Qy 566 ----- 566
Db 4732 AACAGAGAAATTAATTAATTCCTGCTGAGAGAAATATCTGGAGATTATTCAGAAATTTGTA 4791
Qy 567 -----GlnSerThrlValSerLeuGluAerAlaMetProGlu-----AlaLeuLysSer 582
Db 4792 CAAAGGAAACAGATATTAAGAACTGTGATGATATTCAGAAAGATACCAATATTAAGAAA 4851
Qy 583 LeuLeuPheProAerPheGluProLeuThrlValSerPheHleuThrlAerPheLeuLys----- 600
Db 4852 GTAGTTTTTGAAGATAC-----CTCCAGACACATTTTACCAAAACGTGAT 4896
Qy 601 -----GluLeuGluGlnAerGluAlaLeuThrlGluGlyAerSerAenAlaGlnIle 617
Db 4897 CTTGAGAGGAGATGTAAGACATCTCCGAAATTCCTGCTCCTAATATTCAGAGCTGAAGTGT 4956
Qy 618 ArgAerThrlGlnAerGlyLeuValAerValMetLeuLysAenIleGlnIleMetLysHle----- 636
Db 4957 GATGAT-----GTTATCCAAAGTTCGTTTAAATTTAAGCATCCAGACCGGATTTCT 5007
Qy 637 LeuAlaAlaHleuLeuThrlValSerGlnAlaLeuGlnAlaLeuGlnAlaValGlnIleLys 656
Db 5008 CTACATGCCCATGAGATCTTCTTATGAAATATCATCATCAAGAGAAAGATTAAGATGAC 5067
Qy 657 SerSerAerGluGlnAerPheGluAerPheGluAerPheGluGlnValAerGlyLeu 676
Db 5068 TCTCTCT-----GAATGCTTAAAGAAATATGCTGTTCCAG----- 5103
Qy 677 ProLeuAenThrlPheLeuLeuAerGluProLeuThrlValGlyMetHleGlyValGlnValLeu 696
Db 5104 -----CCAAATAGCACTTATACCTTACCTTACCTGACATGCCACT 5139
Qy 697 GluAerGluGlyValAerThrlValSerProSerThrlValAerPheAerGluAerGlyAerAla--- 715
Db 5140 GAGCGATCAGAGGAGAAAGTCTGGGCTCT-----GCCTGTCCGGGCTTGG 5184

```

```

Qy 716 AlaLeuAlaGluIleThrlGluMetValAlaGlnLeuHleGlyThrlMetIleLysMetGlu 735
Db 5185 GCCCTACTACGAGCTGTGAACCCAGAAAAGATATTTCTCAGGCTGTGATAGTCCCTC 5244
Qy 736 AenPheGlnLysLeuHleGluLeuLysLysAerPheLeuIleGlyIleAerAenLeuVal 755
Db 5245 CTATCTGCCAAAAGAAATCTACATTAAGAC-----AGCAACATGCCCTGTG 5292
Qy 756 ProGluAerGluPheIleAerGlySerLeuSerLysLeuSerGlyLysGlyLeuGln 775
Db 5293 GACATGAGAGAAATTTGCTTACTA----- 5316
Qy 776 GlnAerMetPhePheLeuPheAerAerValLeuLeuTy----- 788
Db 5317 -----TTTATGACCTTGTGATGAAGAAAGAGCTGATCTATGAAGAAAGACTCCGA 5367
Qy 789 ThraAerGlyLeuThrlAlaSerAenGlnPheLysValHleGlyGlnLeuProLeuTy 808
Db 5368 AGTCTTGAGACATCACAATCTCCAGAAATGAATAATCCCATGACTTTCAGCCATTAAAT 5427
Qy 809 GlyMetThr-----IleGluLysSerGluAerGluThrlGlyValProHle 823
Db 5428 GCGATGATCTACAGCTTGCCTGCGCTGAATAATGTATAGCAAGAGTGGTGAGTTACAC 5487
Qy 824 CysLeuThrlLeuAerGlyGlnAerGlnSerIleIleValAlaAlaSerSerAerGlu 843
Db 5488 CTGCTGAACATGAGCGCTCCCAAGACATTCAGCTGTGCTTTCACGCGCAACCTTG 5547
Qy 844 MetGluLysThrlValGluAerPheIleGlnMetAlaIleAerPheAlaGluLysSerSer 863
Db 5548 CTGAAAATGACATTAACAGACACAGTGAAGGCTGCGCCCTTTCCTGCTTCAATTT 5607
Qy 864 ProAlaProGluPheLeuAlaSerSerPro 873
Db 5608 AAAACTCTGAATGAAGGATCAAAACCT 5637

RESULT 15
PCT-US06-14800-70
; Sequence 70, Application PC/TUS0614800
; GENERAL INFORMATION:
; APPLICANT: Children's Hospital Medical Center
; APPLICANT: Wong, Hector R.
; APPLICANT: Aronow, Bruce J.
; TITLE OF INVENTION: METALLOTHIONEIN AS AN EARLY BIOMARKER
; TITLE OF INVENTION: FOR DEATH SECONDARY TO SEPTIC SHOCK AND AS A NOVEL
; FILE REFERENCE: CHM027.001VPC
; CURRENT APPLICATION NUMBER: PCT/US06/14800
; PRIOR FILING DATE: 2006-04-19
; PRIOR APPLICATION NUMBER: US 60/673,656
; PRIOR FILING DATE: 2005-04-20
; NUMBER OF SEQ ID NOS: 120
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 70
; LENGTH: 6914
; TYPE: DNA
; ORGANISM: Homo sapiens
PCT-US06-14800-70

Alignment Scores:
Pred. No.: 32.4 Length: 6914
Score: 125.00 Matches: 162
Percent Similarity: 32.2% Conservative: 80
Best Local Similarity: 21.5% Mismatches: 279
Query Match: 2.3% Indels: 232
DB: 1 Gaps: 35

US-09-555-342B-2 (1-1045) x PCT-US06-14800-70 (1-6914)
Qy 285 AenSerAlaTyGlnAerThrlLeuGluPhe-----LeuMetAlaSerAerGluAer 300

```

Db 3690 AACTCAGCCAGCAAAACCTCTCTCCAGACCTCAGCCACAGACCTCTCTCCAGACCTCA 3749
 Qy 301 PheCysIysSerPheTrpValIleCysVal-GluHisIleAlaPhePheArgLeuPheG1 320
 Db 3750 TTCAGAGAAACCTTTCCCGAGCCCTCGGTCAAGTGGCCAT-----TTCCTC 3794
 Qy 320 uGluProIysPro-----LysProIysProVal---LeuPheSerArgG1 334
 Db 3795 CAGACCTCAGCCATCAAAACCTTTCTCCAGACCTCAGCCATCAAAACCTTTCTTAGACC 3854
 Qy 334 ySerSerPheArgPheSerG1ArgThrGlnIysGlnValLeuAspTrpValIysGluG1 354
 Db 3855 TCAGCCAGACAAACCTCTCTCCAGAACTCACTCA-----GACAAACCTTTCTCCAG 3905
 Qy 354 uGlyHisIleIysValGlnPheG1IuArgIysIleSerIysIleHisSerIleArgSerIe 374
 Db 3906 CCGTCGTCAGATGCCCTTTCTCCAGACCTCAGCCATCAAAACCTTTCTTAGA----- 3960
 Qy 374 uAlaSerGlnPro-ThrGluIysLeuAsnSerGluValLeu-----GluG 388
 Db 3961 ----CTTCAGCCAGCAAAACCTCTCCAGAACTCAGCCATCAAGTACTCTCTCCAGAAC 4016
 Qy 388 IAsnSerGlnIysSerThrSerIleuThrPheG1IysGluIysAlaGluSerProG1IysGlns 408
 Db 4017 TCAGTCAGACAAACCTTTCTCCAGCCCTCGGTCAAGTCCCATTTCTCTCCAGACCT-CAGC 4075
 Qy 408 eCyArg-Arg-----ArgG1IysGluProIysValSerAlaG1IysIleProG1IysSerH 425
 Db 4076 CATCAAAACCTTTCTTACACTTCAAGCCAGACAAACCTCTCTCCAGAACTCAAGTCAAAACA 4135
 Qy 425 IAspSerProAlaProArgArg---SerProAlaGluIysIysGlnIleAlaAspIysAla 444
 Db 4136 AACCTTTCCCGAGCCCTCGGTCAAGTGGCCCTTTCTCCAGACCCAGCCATCAAAACCTT 4195
 Qy 444 IAsnSerAlaProThrGlnIysGluIysGlnValValIysAspArgThrGlnIysIysSer 464
 Db 4196 TCTCTAGACCTCAGCCAGCAAAACCT-----CTCTCCAGAACTCAAGTCAAGCAAAAC 4246
 Qy 464 roGlnPro-----Pro-Gln 468
 Db 4247 CTTTCCCAAGACCTCAGTGAATGCCCTTTTCCAGATCTCAATCAATTTCCCTTACC 4306
 Qy 469 ProSerThrIysLeuThrGlySerProHisLeuSerGlu-----LeuSerValAsn 486
 Db 4307 CCAAGCCTCGACAGATGACACTTCTCCAGACCTCGGTGAGACAGATCTTTCCCAAAAC 4366
 Qy 487 ---SerGlnIysGluValAlaPro-----AlaAsnValThrLeuSerProAsnLeuSer 503
 Db 4367 TTGGTCAAGATGTCCTTTCCCGAGACCTCAGCAGGTGACTCTCTCCAGACATCAGT 4426
 Qy 504 -----ProSerThrIysGlnIleAsnSerPro----- 511
 Db 4427 GACACCAACCTTCTCCGGATCTCAGCCAGATACACTCTCTCCAGACCTTGATCAATA 4486
 Qy 512 -----LeuIle----- 513
 Db 4487 TTTCACCTTCTGAATCTAGTCAATCATCTCTTCAAGAAATTAAATGAGTCTTTCC 4546
 Qy 514 -----SerProLeuLeuAsnPro----- 519
 Db 4547 TATTCAGACCTTGGTCAGATGCCATCTCTTCATCTCTCACTCAATGATATCTTTCTA 4606
 Qy 519 ----- 519
 Db 4607 TCAAAGAAATTTAATCCACTGTTATAGTGGCTCAGTAAAGTGTACAGATTACATT 4666
 Qy 520 GlnAlaCysProArg-----ThrAspAspGluAspGluIysArgIysValArg 535
 Db 4667 GAGATCATTCCTCAAGGAAGAGGTCCAGAGCAGTGAAGTACTATGCTGAATGATGATTAT 4726
 Qy 536 PheProThrAspIysValIysPheIleAlaIysGluValSerThrThrGluAspThrTyr 555
 Db 4727 GTGCCCTATGATGACCCCTTAC-----AAACTGATGTTAAGAGCAAAAC 4768

Qy 556 Leu-----LysAspLeuGluValIleThrSerTrpPhe----- 566
 Db 4769 ATCACTCTCCAGAGATCTCTGACAAATTCGACATGTGATACCTCCGACAGCAAAATGGA 4828
 Qy 566 ----- 566
 Db 4829 AACAGAAATTTAATCAATGCTGCTGAAGAAATATCTGGGATTAATGCAATTTGTA 4888
 Qy 567 -----GlnSerThrValSerIysGluAspAlaMetProGlu-----AlaLeuIysSer 582
 Db 4889 CAAAGGAAACAAATATTAAGACCTGTGATGATATTCAGAAAGATACACATTTAAAGAA 4948
 Qy 583 LeuIlePheProAsnPheGluProLeuHisIleIysPheHisIleThrAsnPheLeuIys----- 600
 Db 4949 GTAGTTTTCGAAGTAC-----CTGACAGACACTTTTACCAAAACGTGAT 4993
 Qy 601 -----GluIleGluIysValLeuAlaLeuTrpGluIysIysSerValAlaGlnIle 617
 Db 4994 CTTGAGGAGATGATGAAGACATCTCGAAATCTTGGTCTTATTAACAGAGCTGAAGTG 5053
 Qy 618 ArgAspTrpGlnArgIleGluAspValMetLeuIysAsnIleGlnIysMetIysVal--- 636
 Db 5054 GATGAT-----GTTATCCAAAGTTCGTTTAAATTTAAGATCCAGACCGTATTCT 5104
 Qy 637 LeuAlaAlaIysIleuTrpIysIysSerGluAlaLeuGluIysIysValIleIys 656
 Db 5105 CTACATGCCCATGAGCTTCTCTGAAATAATCAAGAGGAAAGACTTATGAAGATGAC 5164
 Qy 657 SerSerArgArgLeuGluAsnPheCysValAspPheGluLeuGlnIysValCysValTrpLeu 676
 Db 5165 TCTCTCT-----GATAGTTTTAAAGAAATATGCTGTAG----- 5200
 Qy 677 ProLeuAsnThrPheLeuIysGluIysValIysGluIysValIleIys 696
 Db 5201 -----CDAATGAGATTAATCTTACGATGAGCTGACCT 5236
 Qy 697 GluArgLeuCysValHisIysProProSerHisValAspPheArgAspCysValGluAla--- 715
 Db 5237 GAGCGATACAGGCGCAGAAAGTCTGGCTCT-----GCTGTCCGGCTTGG 5281
 Qy 716 AlaLeuAlaGluIleThrGluMetValAlaGlnLeuHisIysIleThrMetIleIysMetGlu 735
 Db 5282 GCGTACTACTCAGCTGTGAACCCAGAAAGAAATATCACTCAGCGTGAATGAGTCCCTC 5341
 Qy 736 AsnPheGlnIysLeuHisIysGluIysValIysAspLeuIleGluIleAspAsnLeuValI 755
 Db 5342 CTATCTGCCAAAGAAAGATATCAATAAAGAC-----AGCAACATGCTGTG 5389
 Qy 756 ProGluArgGluPheIleArgLeuGluIysLeuSerIysLeuSerGluIysGluLeuGln 775
 Db 5390 GACATGAGAAATTTGCTTACTA----- 5413
 Qy 776 GlnArgMetPhePheLeuPheAsnAspValLeuLeuTrp----- 788
 Db 5414 -----TTTATGACCTTGTATGAAGAAAGAGCTGATATGAAAGAACTCCGA 5464
 Qy 789 ThrSerArgGluLeuThrAlaSerAsnGlnPheIysValHisIysGlnIleuProLeuTrp 808
 Db 5465 AGTTCCTGAGACTACATCTCAGAAATGAAGAAATCCATGAGTTTCAAGCCATTAAT 5524
 Qy 809 GlyMetThr-----IleGluIysSerGluAspGluIysTrpGluValProHis 823
 Db 5525 GGGATGATCTACAGCTTGGCTGCGCTGAAATATGATAGCAAGAGTGGAGGTTCAC 5584
 Qy 824 CysLeuThrIysArgGluIysIysGlnIysIleIleValAlaIleAsnSerArgSerGlu 843
 Db 5585 CTGCTGAACATACAGCGCTCCCAAGCAATTCAGGTGATTCACCTTCAAGCCAGACCTTG 5644
 Qy 844 MetGluIysTrpValGluAspIleGluMetAlaIleAspLeuAlaGluIysSerSerSer 863
 Db 5645 CTGAAATAGGCAATAAACAGCACACAGTTAAGGGGTCTGGCCCTTCTGCTGATTCATT 5704

Oy 864 ProAlaProGluPheLeuAlaSerSerPro 873
|||
Db 5705 AAACCTTGGAAATGAGGACATCAAACT 5734

Search completed: May 20, 2006, 06:59:15
Job time : 795 secs

This Page Blank (uspto)